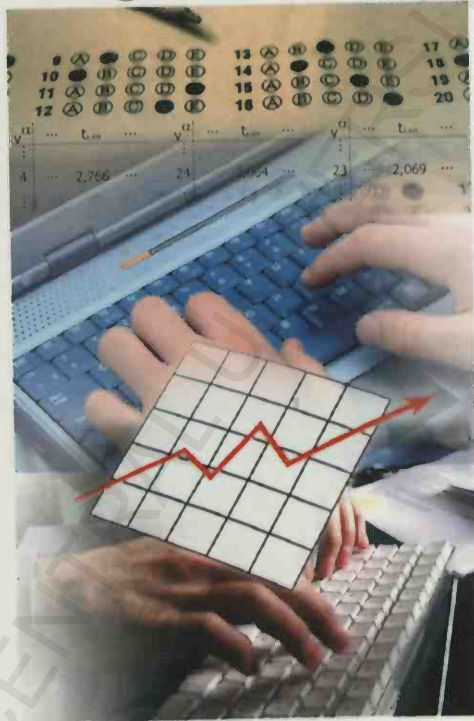




STATISTICĂ

teste grilă și probleme



ELISABETA JABA

CARMEN PINTILESCU

editura
sed
com
libris

III 31.851

BIBL. CENTR. UNIV.
„M. EMINESCU” IAȘI

III 31.851

STATISTICĂ

teste grilă și probleme

teste grilă și probleme

Descrierea CIP a Bibliotecii Naționale a României
JABA, ELISABETA
Statistici: teste grilă și probleme / Elisabeta Jaba.
Gărnă Piniștea. - Ed. a 2-a. - Iași: Sedcom Libris, 2007
Bibliogr.
I.S.B.N.: 978-973-870-527-2
I. Piniștea, Gărnă
911(075.8)(079.1)

Ediția Sedcom Libris este acreditată de Comisia Națională de Cercetări
Științifice din Învățământul Superior (C.N.C.S.I.S.)

Copyright © 2007, SEDCOM LIBRIS

Toate drepturile asupra prezentei ediții sunt rezervate Editurii
Sedcom Libris. În caz de reproducere, distribuție sau utilizare a
textelor prin orice mijloc, trebuie să se obțină din partea Editurii
Sedcom Libris, este interzisă și se va pedepsi
conform legislației în vigoare.

Adresa: 202, Măia de Foc nr. 4, Gărnă Piniștea
Tel.: 0232.242.877; 234.582; fax: 0232.234.580

e-mail: editionsedcomlibris@yahoo.com www.sedcom.ro

Referenți științifici:

prof. univ. dr. Elena Maria BIJI

prof. univ. dr. Ana GRAMA

Redactor: Petru RADU

Revizia textelor: Alina HUCAI

Concepția și realizarea copertei: Liliana BASARAB

Prepress: Ciprian TOȘCARIU

Tehnoredactarea aparține autoarelor.

Descrierea CIP a Bibliotecii Naționale a României

JABA, ELISABETA

Statistică: teste grilă și probleme / Elisabeta Jaba,
Carmen Pintilescu. – ed. a 2-a rev. – Iași: Sedcom Libris, 2007
Bibliogr.

I.S.B.N.: 978-973-670-227-3

I. Pintilescu, Carmen

311(075.8)(079.1)

Editura Sedcom Libris este acreditată de Consiliul Național al Cercetării Științifice din Învățământul Superior (C.N.C.S.I.S.).

Copyright © 2007, SEDCOM LIBRIS

Toate drepturile asupra prezentei ediții sunt rezervate Editurii Sedcom Libris, Iași. Reproducerea parțială sau integrală a textelor, prin orice mijloc, precum și a graficii copertei, fără acordul scris al Editurii Sedcom Libris, este interzisă și se va pedepsi conform legislației în vigoare.

Adresa: Șos. Moara de Foc nr. 4, Iași, România

Tel.: 0232.242.877; 234.582; fax: 0232.233.080

e-mail: editurasedcomlibris@yahoo.com; www.sedcom.ro

ELISABETA JABA

CARMEN PINTILESCU

STATISTICĂ

teste grilă și probleme

455 133

ediția a II-a revăzută



444782

B.C.U. IASI

ALE

editura



libris

Elisabeta JABA (n. 1946) este profesor doctor la Universitatea „Al. I. Cuza” Iași, șeful Catedrei de Statistică și Prognoză, conducător de doctorat în specializarea Statistică economică. A absolvit Facultatea de Științe Economice din Iași, în 1969 și a obținut titlul științific de doctor în economie, specializarea Statistică economică, în 1979. A efectuat mai multe stagii de perfecționare și schimburi de experiență la universități din Franța (în Gestiunea resurselor umane, Anchete prin sondaj, 1992, 1997, 2000), Spania (Statistică și econometrie, 1992), Anglia (Anchete prin sondaj, 2002), Republica Moldova (1997-2004), Italia (2003). A inițiat și coordonat programe SOCRATES și ERASMUS, în cooperare cu universități din Franța, Italia, Republica Moldova. A publicat, ca unic autor sau coautor, 26 cărți și cursuri universitare, precum și 89 articole științifice, în țară și străinătate (China, Franța, Belgia, Italia). Pentru cartea „Statistică”, Ediția a treia (Editura Economica, București, 2002) a primit Premiul anului 2002 pentru cel mai bun manual al anului în domeniul științei economice fundamentale, conferit de AGER – Asociația Generală a Economistilor din România, precum și Premiul „Cel mai elevat curs universitar de statistică” acordat de Societatea Română de Statistică, București, 23 decembrie 2004.

Este titularul cursurilor „Statistică”, „Teoria și practica sondajului statistic”, „Software statistic”. Face parte din Colegiul științific al Revistei Române de Statistică, este expert evaluator al CNCIS, precum și expert în Eșantionare, confirmat de Department for International Development (DFID) al Guvernului Marii Britanii, pentru proiectul ISAS. Este președintele Societății Române de Statistică – Filiala Iași și membru al mai multor societăți profesionale și științifice naționale și internaționale de prestigiu, printre care A.G.E.R. și Societatea Franceză de Statistică.

Carmen Pintilescu este conferențiar universitar doctor la Universitatea „Al. I. Cuza” Iași, Facultatea de Economie și Administrarea Afacerilor, Catedra de Statistică și Prognoză. A absolvit Facultatea de Științe Economice din Iași în anul 1992. A obținut Diploma de Studii Aprofundate, în anul 1996, la Universitatea din Poitiers (Franța), iar în perioada 1996-2001 a fost doctorand al aceleiași universități. În anul 2001 a obținut titlul științific de doctor în economie la Universitatea din Poitiers (Franța). Este autor sau coautor a 7 cărți și cursuri universitare și a publicat 30 articole și studii din domeniul statisticii. Este titular al cursului de Analiza datelor.

CUPRINS

PREFAȚĂ	11
Capitolul 1 ROLUL STATISTICII ÎN CUNOAȘTERE.....	13
1.1. CE ESTE STATISTICA ?	14
1.2. PARTICULARITĂȚI ȘI ETAPE ALE PROCESULUI CUNOAȘTERII STATISTICE.....	14
1.3. NOȚIUNI FUNDAMENTALE ALE STATISTICII	15
1.4. ALEGEȚI RĂSPUNSUL CORECT	17
1.5. RECUNOAȘTEȚI EXEMPLUL CORECT.....	25
1.6. RĂSPUNSURI CORECTE.....	27
Capitolul 2 OBSERVAREA STATISTICĂ.....	33
2.1. PRINCIPII DE BAZĂ ALE OBSERVĂRII STATISTICE.....	34
2.2. SURSE DE DATE STATISTICE	34
2.3. PROGRAMUL UNEI OBSERVĂRI STATISTICE	34
2.4. METODE DE ÎNREGISTRARE STATISTICĂ	35
2.5. LUCRĂRI DE ÎNREGISTRARE STATISTICĂ	35
2.6. ERORI DE ÎNREGISTRARE STATISTICĂ.....	36
2.7. ALEGEȚI RĂSPUNSUL CORECT.....	37
2.8. RECUNOAȘTEȚI EXEMPLUL CORECT.....	43
2.9. RĂSPUNSURI CORECTE.....	44
Capitolul 3 SISTEMATIZAREA ȘI PREZENTAREA DATELOR STATISTICE	49
3.1. SISTEMATIZAREA DATELOR	50
3.2. PREZENTAREA DATELOR STATISTICE	50
3.3. PREZENTAREA DISTRIBUȚIILOR STATISTICE UNIDIMENSIONALE.....	51
3.4. PREZENTAREA DISTRIBUȚIILOR STATISTICE BIDIMENSIONALE.....	52
3.5. REPREZENTAREA GRAFICĂ A UNEI DISTRIBUȚII ÎN SPSS	53
3.6. ALEGEȚI RĂSPUNSUL CORECT.....	56
3.7. RECUNOAȘTEȚI EXEMPLUL CORECT.....	65
3.8. RĂSPUNSURI CORECTE.....	66
3.9. PROBLEME.....	71
Capitolul 4 INDICATORI STATISTICI ÎN MĂRIMI ABSOLUTE ȘI ÎN MĂRIMI RELATIVE	77
4.1. INDICATORI ÎN MĂRIMI ABSOLUTE	78
4.2. INDICATORI ÎN MĂRIMI RELATIVE	78
4.3. CALCULUL MĂRIMILOR RELATIVE DE STRUCTURĂ ÎN SPSS.....	79
4.4. ALEGEȚI RĂSPUNSUL CORECT.....	80
4.5. RECUNOAȘTEȚI EXEMPLUL CORECT.....	84
4.6. RĂSPUNSURI CORECTE.....	85
4.7. PROBLEME.....	88
Capitolul 5 INDICATORI AI TENDINȚEI CENTRALE.....	91
5.1. PROBLEME GENERALE ALE MĂRIMILOR MEDII	92
5.2. MEDIA ARITMETICĂ	92
5.3. MEDII CU APLICAȚIE SPECIALĂ	94

5.4. MODUL (M_0) SAU DOMINANTA (D_0).....	94
5.5. MEDIANA (ME).....	96
5.6. MEDIALA (ML)	98
5.7. RELAȚII ÎNTRE VALORILE TENDINȚEI CENTRALE.....	99
5.8. ALEGEREA CELEI MAI SEMNIFICATIVE VALORI CENTRALE	99
5.9. CALCULUL INDICATORILOR TENDINȚEI CENTRALE ÎN SPSS.....	99
5.10. ALEGEȚI RĂSPUNSUL CORECT.....	101
5.11. RECUNOAȘTEȚI EXEMPLUL CORECT.....	109
5.12. RĂSPUNSURI CORECTE.....	111
5.13. PROBLEME.....	116
Capitolul 6 INDICATORI AI DISPERSIEI, ASIMETRIEI ȘI BOLTIRII	129
6.1. INDICATORI SIMPLI ȘI SINTETICI AI DISPERSIEI	130
6.2. PROPRIETĂȚI ȘI CALCULUL SIMPLIFICAT AL VARIANȚEI.....	130
6.3. INDICATORI AI DISPERSIEI UNEI VARIABILE ALTERNATIVE.....	131
6.4. MĂSURAREA DISPERSIEI ÎN SISTEMUL MEDIANEI	131
6.5. INDICATORI FACTORIALI AI DISPERSIEI	132
6.6. INDICATORI AI FORMEI.....	133
6.7. CALCULUL INDICATORILOR DISPERSIEI, ASIMETRIEI ȘI BOLTIRII ÎN SPSS	134
6.8. ALEGEȚI RĂSPUNSUL CORECT.....	136
6.9. RECUNOAȘTEȚI EXEMPLUL CORECT.....	146
6.10. RĂSPUNSURI CORECTE.....	153
6.11. PROBLEME.....	165
Capitolul 7 INDICATORI AI CONCENTRĂRII.....	175
7.1. NOȚIUNI.....	176
7.2. PROCEDEE GRAFICE DE DETERMINARE A CONCENTRĂRII (CURBA DE CONCENTRARE) ...	176
7.3. PROCEDEE NUMERICE DE DETERMINARE A CONCENTRĂRII	177
7.4. ALEGEȚI RĂSPUNSUL CORECT.....	178
7.5. RECUNOAȘTEȚI EXEMPLUL CORECT.....	182
7.6. RĂSPUNSURI CORECTE.....	183
7.7. PROBLEME.....	186
Capitolul 8 INDICATORI AI SERIILOR CRONOLOGICE.....	195
8.1. INDICATORI PENTRU CARACTERIZAREA NIVELULUI ȘI VARIAȚIEI ÎN TIMP	196
8.2. ÎNDICII DINAMICII	196
8.3. CALCULUL INDICATORILOR SERIILOR CRONOLOGICE ÎN SPSS	197
8.4. ALEGEȚI RĂSPUNSUL CORECT.....	199
8.5. RECUNOAȘTEȚI EXEMPLUL CORECT.....	204
8.6. RĂSPUNSURI CORECTE.....	207
8.7. PROBLEME.....	211
Capitolul 9 EȘANTIONAREA	225
9.1. CE ESTE CERCETAREA PRIN SONDAJ ?	226
9.2. NOȚIUNI ȘI TERMENI PERECHI	226
9.3. DEMERSUL CERCETĂRII PRIN SONDAJ	228
9.4. PROBLEME ALE EȘANTIONĂRII.....	228
9.5. ERORI DE REPREZENTATIVITATE.....	229

9.6. CALCULUL MĂRIMII ERORILOR DE REPREZENTATIVITATE.....	229
9.7. ALEGEȚI RĂSPUNSUL CORECT.....	230
9.8. RĂSPUNSURI CORECTE.....	232
9.9. PROBLEME.....	233
9.9.1. Determinarea volumului eșantionului.....	233
9.9.2. Repartizarea volumului eșantionului pe grupe tipice.....	234
9.9.3. Determinarea numărului de eșantioane posibil de extras dintr-o populație.....	235
Capitolul 10 ESTIMAREA PARAMETRILOR UNEI POPULAȚII.....	237
10.1. PROBLEME GENERALE ALE UNEI ESTIMĂRI.....	238
10.2. PROPRIETĂȚI ALE ESTIMATORILOR.....	239
10.3. ESTIMAREA PUNCTUALĂ A PARAMETRILOR UNEI POPULAȚII.....	239
10.3.1. Estimatorul $\hat{\mu}$ al mediei μ a unei populații.....	240
10.3.2. Estimatorul $\hat{p}$ al proporției p a unei populații.....	241
10.3.3. Estimatorul $\hat{\sigma}^2$ al varianței σ^2 a unei populații.....	241
10.3.4. Estimatorul varianței distribuției de selecție a diferenței dintre două medii.....	242
10.3.5. Estimatorul varianței distribuției de selecție a diferenței dintre două proporții.....	243
10.4. ESTIMAREA PRIN INTERVAL DE ÎNCREDERE A PARAMETRILOR UNEI POPULAȚII.....	243
10.4.1. Definirea intervalului de încredere (I.C.).....	243
10.4.2. Estimarea mediei prin interval de încredere.....	244
10.4.3. Estimarea proporției prin interval de încredere.....	245
10.5. ESTIMAREA PARAMETRILOR UNEI POPULAȚII CU SPSS.....	245
10.5.1. Estimarea mediei μ a unei populații.....	245
10.5.2. Estimarea proporției p a unei populații.....	246
10.6. ALEGEȚI RĂSPUNSUL CORECT.....	247
10.7. RECUNOAȘTEȚI EXEMPLUL CORECT.....	251
10.8. RĂSPUNSURI CORECTE.....	255
10.9. PROBLEME.....	261
10.9.1. Calculul probabilității pentru o distribuție statistică.....	261
10.9.2. Estimarea erorii medii de selecție a mediei.....	263
10.9.3. Estimarea erorii limită a mediei.....	264
10.9.4. Estimarea mediei prin interval de încredere.....	266
10.9.5. Estimarea proporției prin interval de încredere.....	273
10.9.6. Estimarea prin interval de încredere a diferenței dintre două medii.....	277
10.9.7. Estimarea prin interval de încredere a diferenței dintre două proporții.....	279
Capitolul 11 TESTAREA IPOTEZELOR STATISTICE.....	281
11.1. DEMERSUL TESTĂRII UNEI IPOTEZE STATISTICE.....	282
11.2. TESTAREA IPOTEZELOR STATISTICE ASUPRA UNUI PARAMETRU.....	284
11.3. VERIFICAREA NORMALITĂȚII UNEI DISTRIBUȚII.....	286
11.4. TESTAREA IPOTEZELOR STATISTICE CU SPSS.....	287
11.4.1. Testarea egalității unei medii cu o valoare specificată.....	287
11.4.2. Testarea egalității mediilor a două eșantioane independente.....	287
11.4.3. Testarea egalității mediilor a două eșantioane perechi.....	288
11.4.4. Testarea egalității unei proporții cu o valoare specificată.....	289

11.4.5. Testarea egalității a două și mai multe proporții.....	289
11.5. ALEGEȚI RĂSPUNSUL CORECT.....	291
11.6. RECUNOAȘTEȚI EXEMPLUL CORECT.....	295
11.7. RĂSPUNSURI CORECTE.....	301
11.8. PROBLEME.....	307
11.8.1. Testarea semnificației mediei unei populații.....	307
11.8.2. Testarea semnificației unei proporții.....	309
11.8.3. Testarea ipotezelor privind compararea a două medii.....	311
11.8.4. Testarea normalității unei distribuții.....	313
Capitolul 12 ANALIZA DISPERSIONALĂ (ANOVA).....	319
12.1. COMPONENTE ALE VARIAȚIEI. TIPURI DE ANOVA.....	320
12.2. COMPONENTE ALE VARIAȚIEI ȘI ESTIMATORI AI VARIANȚELOR ÎN ANOVA UNIFACTORIALĂ.....	321
12.3. CONDIȚII. IPOTEZE. REGULA DE DECIZIE ÎN ANOVA UNIFACTORIALĂ.....	322
12.4. ANOVA CU SPSS.....	323
12.5. ALEGEȚI RĂSPUNSUL CORECT.....	325
12.6. RECUNOAȘTEȚI EXEMPLUL CORECT.....	328
12.7. RĂSPUNSURI CORECTE.....	330
12.8. PROBLEME.....	332
Capitolul 13 ANALIZA DE REGRESIE ȘI CORELAȚIE.....	337
13.1. PROBLEME ALE ANALIZEI DE REGRESIE ȘI CORELAȚIE.....	338
13.2. METODE DE ANALIZĂ A LEGĂTURILOR STATISTICE.....	339
13.3. MODELE DE REGRESIE CU DOUĂ VARIABILE.....	339
13.3.1. Estimarea punctuală a parametrilor ecuației de regresie.....	340
13.3.2. Estimarea prin interval de încredere.....	340
13.4. MĂSURAREA INTENSITĂȚII LEGĂTURII DINTRE DOUĂ VARIABILE.....	341
13.4.1. Coeficientul de corelație.....	341
13.4.2. Raportul de corelație.....	342
13.5. CORELAȚIA NEPARAMETRICĂ.....	344
13.6. ANALIZA DE REGRESIE ȘI CORELAȚIE CU SPSS.....	344
13.7. ALEGEȚI RĂSPUNSUL CORECT.....	347
13.8. RECUNOAȘTEȚI EXEMPLUL CORECT.....	352
13.9. RĂSPUNSURI CORECTE.....	356
13.10. PROBLEME.....	361
13.10.1. Estimarea parametrilor modelului de regresie.....	361
13.10.2. Testarea semnificației coeficientului de regresie β	365
13.10.3. Analiza de corelație.....	367
13.10.4. Corelația neparametrică.....	371
Capitolul 14 ANALIZA ȘI EXTRAPOLAREA SERIILOR CRONOLOGICE.....	377
14.1. COMPONENTELE UNEI SERII CRONOLOGICE.....	378
14.1.1. Componente ale unei serii cronologice.....	378
14.1.2. Combinarea componentelor unei serii cronologice.....	378
14.2. AJUSTAREA SERIILOR CRONOLOGICE.....	379
14.3. AJUSTAREA SERIILOR CRONOLOGICE SEZONIÈRE.....	381

14.4. CRITERII DE ALEGERE A CELUI MAI BUN MODEL DE TREND	381
14.5. PROGNOZA STATISTICĂ	382
14.5.1. Probleme de prognoză	382
14.5.2. Orizontul prognozei	383
14.5.3. Extrapolarea nivelului unui fenomen pentru un orizont de prognoză dat	383
14.5.4. Determinarea zonei de prognoză	384
14.6. ALEGEȚI RĂSPUNSUL CORECT	385
14.7. RECUNOAȘTEȚI EXEMPLUL CORECT	389
14.8. RĂSPUNSURI CORECTE	391
14.9. PROBLEME	394
14.9.1. Ajustarea seriilor de timp. Modelul de trend liniar	394
14.9.2. Prognoza statistică	396
14.9.3. Ajustarea seriilor de timp cu caracter sezonier	399
ANEXE	405
BIBLIOGRAFIE	419

modelarea lor, cât și aplicarea acestor modele. Cu ajutorul metodelor se obține informația necesară pentru luarea deciziilor și se elaborează previziuni în funcție de variația fenomenelor. În concluzie, toate rîc asumate.

Demersul de lucru se desfășoară pe un raționament propriu, indicând deductiv înaintea planului de lucru, de obicei, și ajungând la obiectul de studiu, pentru a realiza o lucrare de cercetare.

Lucrarea este structurată în capitole în funcție de conținutul și de scopul din anul I și II de studiu. În primul capitol se prezintă scopul și obiectul de studiu, iar în al doilea capitol se prezintă obiectul de studiu și obiectul de studiu. În al treilea capitol se prezintă obiectul de studiu și obiectul de studiu.

„Statistica” este o disciplină care se ocupă de studiul și analiza datelor. Fiecare capitol din lucrare este structurat în secțiuni. Prima secțiune „Alegeți răspunsul corect” este o secțiune de tip test, care are ca scop evaluarea cunoștințelor de la începutul cursului. A doua secțiune „Recunoașteți exemplul corect” permite aprofundarea elementelor principale ale disciplinei. Secțiunea a treia „Răspunsuri corecte” permite autoevaluarea, precum și verificarea rezultatelor din primele două secțiuni. A patra secțiune „Probleme” permite aprofundarea tehnicilor de calcul statistic și a modului de interpretare a rezultatelor și a fenomenelor.

Structurată ca un instrument complementar, lucrarea ajută la fixarea și verificarea cunoștințelor preluate la cursul de Statistică, potând descriptivă și inferențială.

Ca obiectiv, își propune să faciliteze înțelegerea noțiunilor esențiale ale acestei discipline științifice și însușirea acestor noțiuni într-un mod clar și

PREFAȚĂ

Statistica a devenit elementul sine qua non al cunoașterii în condiții de incertitudine. Utilizarea statisticii permite atât descrierea fenomenelor și modelarea lor, cât și aplicarea practică a acestora. Cu ajutorul statisticii se obține informația necesară fundamentării deciziei și se elaborează previziuni în funcție de variația fenomenului studiat și în condițiile unui risc asumat.

Demersul statistic se bazează pe un raționament propriu, inductiv-deductiv iterativ, plecând de la concretul datelor și ajungând la abstractul unui model, pentru a reveni la concretul previziunii.

Lucrarea este destinată studenților în științe economice, în special celor din anul I și II, și are ca scop inițierea acestora în demersul statistic și formarea gândirii statistice necesare înțelegerii fenomenelor de masă, în variația lor curentă-continuă în timp, spațiu și din punct de vedere calitativ.

„Statistică. Teste grilă și probleme” este structurată pe 14 capitole, fiecare capitol divizându-se în patru secțiuni. Prima secțiune „Alegeți răspunsul corect” permite să se verifice însușirea corectă a subiectelor tratate. A doua secțiune „Recunoașteți exemplul corect” permite aprofundarea elementelor particular dificile. Secțiunea a treia „Răspunsuri corecte” permite autoevaluarea, prezentând răspunsurile pentru testele din primele două secțiuni. A patra secțiune „Probleme” permite deprinderea tehnicilor de calcul statistic și a modului de interpretare statistică a fenomenelor.

Proiectată ca un instrument complementar, lucrarea ajută la fixarea și verificarea cunoștințelor predate la cursul de Statistică¹, partea descriptivă și inferențială.

Ca obiectiv, își propune să faciliteze asimilarea noțiunilor esențiale ale acestei discipline științifice și însușirea artei de a descrie într-un mod clar și

¹ Elisabeta Jaba, *Statistică*, Ediția a treia, Editura Economică, București, 2002.

concis o colectivitate, folosind informația din datele obținute prin observări statistice asupra fenomenelor studiate.

În prima parte, lucrarea vizează probleme ale statisticii descriptive, iar în a doua parte, probleme ale statisticii inferențiale, și anume: sistematizarea și prezentarea datelor în serii, tabele și grafice, sinteza acestora într-un număr mic de parametri cheie (indicatori ai tendinței centrale, indicatori ai dispersiei, indicatori ai asimetriei, boltirii și concentrării, indicatori ai covariației, indicatori ai seriilor cronologice), eșantionarea, estimarea parametrilor unei populații, testarea ipotezelor statistice, analiza dispersională (ANOVA), analiza de regresie și corelație, analiza și extrapolarea seriilor de timp.

În speranța că lucrarea va fi utilă, mulțumim celor care vor veni cu sugestii și observații atât de necesare îmbunătățirii oricărui lucru supus perfecțării.

Elisabeta JABA,
Martie 2005, Iași

Capitolul 1

ROLUL STATISTICII ÎN CUNOAȘTERE

- ◆ Ce este statistica?
- ◆ Particularități și etape ale procesului cunoașterii statistice
- ◆ Noțiuni fundamentale ale statisticii
- ◆ Alegeți răspunsul corect
- ◆ Recunoașteți exemplul corect
- ◆ Răspunsuri corecte

1.1. Ce este statistica?

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 1-13)

- **Statistica – știință**

- *Obiectul de studiu al statisticii* îl constituie variația curentă-continuu, în timp, spațiu și din punct de vedere calitativ, a fenomenelor de tip stochastic din orice domeniu al vieții economico-sociale sau naturale.

- *Metoda statistică* este definită de un ansamblu de principii metodologice, procedee și tehnici care permit producerea informației statistice, pe baza observării statistice, a prelucrării și analizei datelor statistice, precum și fundamentarea deciziilor privind starea și variabilitatea colectivităților statistice, în timp, spațiu și din punct de vedere calitativ.

- *Scopul statisticii* este cunoașterea, atât aplicativă, cât și fundamentală, a fenomenelor de masă, caracterizate prin variabilitate și produse sub semnul incertitudinii.

- **Statistica – metodă**

- **Statistica – activitate practică**

- **Scurt istoric al dezvoltării statisticii**

- *Prestatistica*

- *Etapa entuziasmului statistic*: școala descriptivă germană și școala engleză a Aritmeticii politice

- *Influența calculului probabilităților*

- *Etapa modernă*

1.2. Particularități și etape ale procesului cunoașterii statistice

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 13-16)

- **Particularități ale raționamentului statistic**
 - *Raționamentul deductiv*
 - *Raționamentul inductiv*
 - *Raționamentul inductiv-deductiv iterativ*
- **Principii metodologice ale statisticii**
 - *Observarea faptică*
 - *Exprimarea numerică*
 - *Comparabilitatea datelor*
- **Etapele demersului statistic**
 - *Definirea problemei*
 - *Observarea statistică*
 - *Prelucrarea statistică*
 - *Analiza statistică*
 - *Decizia statistică*
- **Procedee specifice metodei statisticii**

1.3. Noțiuni fundamentale ale statisticii

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 16-28)

- **Colectivități statistice**
 - O colectivitate statistică reprezintă o mulțime de elemente unite între ele printr-o trăsătură esențial comună, numită omogenitate. Colectivitățile statistice sunt cunoscute în literatura de specialitate sub denumirea de mulțimi, populații.
 - *Clasificarea colectivităților statistice:* colectivități de stări și colectivități de mișcări; colectivități totale și colectivități parțiale.
- **Unități statistice**
 - Unitățile statistice sunt elementele componente ale unei colectivități statistice.

- *Clasificarea unităților statistice:* unități statistice simple și unități statistice complexe; unități statistice active și unități statistice pasive.

- **Variabile statistice**

- Variabilele statistice reprezintă *însușiri, trăsături esențiale purtate de unitățile statistice ale unei colectivități*. Sunt cunoscute în literatură și sub denumirea de *caracteristici statistice*.

- *Simbolizare.* O variabilă statistică se notează cu literă majusculă, de exemplu, X . Este o funcție cu valori reale definită pe mulțimea de bază Ω : $X: \Omega \rightarrow R$. O variantă purtată de o unitate statistică sau un grup de unități, n_i , se notează cu minusculă, de exemplu, x_i .

- *Clasificarea variabilelor statistice:* variabile statistice esențiale și variabile neesențiale; variabile calitative, de timp și de spațiu; variabile numerice (variabile discrete și variabile continue) și variabile nenumerice (nominale și nominal ordinale); variabile alternative și variabile nealternative.

- **Cuantificarea fenomenelor. Tipuri de scală**

- *Scala nominală*
- *Scala ordinală*
- *Scala interval*
- *Scala raport*
- *Scala de intensitate*
- *Scala de opinie*
- *Scala rating*
- *Scala Likert*

- **Indicatori statistici**

- Un indicator statistic poate fi definit ca *rezultatul numeric al unei numărări, al unei măsurări statistice a fenomenelor și proceselor de masă sau al unui calcul asupra datelor obținute printr-o înregistrare statistică*.

- *Clasificarea indicatorilor statistici:* indicatori primari și indicatori derivați; indicatori sintetici și indicatori analitici.



1.4. Alegeți răspunsul corect

1. Crearea cuvântului „statistică“ este atribuită lui
 - a) Gottfried Achenwall
 - b) Dionisie Pop Marțian
 - c) John Graunt
 - d) Edmund Halley
2. De peste șase milenii, statistica este atestată ca
 - a) știință a descrierii statului cu ajutorul cifrelor
 - b) activitate de strângere a datelor, cu caracter militar și fiscal
 - c) știință a gestionării resurselor umane și materiale
 - d) metoda de estimare a speranței medii de viață
3. În etapa „prestatistică“, statistica este cunoscută sub formă de
 - a) numărătoare a populației
 - b) evidență a bunurilor materiale
 - c) instituție a statului
 - d) știință
4. Instituționalizarea statisticii în țara noastră a avut loc
 - a) în 1859, sub conducerea domnitorului Al.I. Cuza
 - b) în 1700, sub conducerea lui Antioh Cantemir
 - c) în 1849, sub conducerea prințului N. Șuțu
 - d) în 1860, sub conducerea lui M. Kogălniceanu
5. Ioan Ionescu de la Brad a scris
 - a) *Notions statistiques sur la Moldavie* (1849)
 - b) *Teoria statisticii sociale* (1972)
 - c) *Povătuiri pentru facerea catagrafiei Moldovei, precedate de oarecare elemente de statistică* (1859)
 - d) *Political Arithmetic* (1676)
6. Alexandru Bărbat a scris
 - a) *Notions statistiques sur la Moldavie* (1849)
 - b) *Teoria statisticii sociale* (1972)
 - c) *Povătuiri pentru facerea catagrafiei Moldovei, precedate de oarecare elemente de statistică* (1859)

- d) *Political Arithmetic* (1676)
7. William Petty a scris
- a) *Notions statistiques sur la Moldavie* (1849)
 - b) *Teoria statisticii sociale* (1972)
 - c) *Povățuiri pentru facerea catagrafiei Moldovei, precedate de oarecare elemente de statistică* (1859)
 - d) *Political Arithmetic* (1676)
8. Prin activitatea practică și opera lor, sunt considerați întemeietori ai statisticii în țara noastră
- a) Ioan Ionescu de la Brad
 - b) Dionisie Pop Marțian
 - c) Herman Corning
 - d) Nicolae Șuțu
 - e) George Barițiu
9. Sunt considerați reprezentanți ai „Aritmeticii politice“
- a) Ioan Ionescu de la Brad
 - b) Edmund Halley
 - c) Dionisie Pop Marțian
 - d) Nicolae Șuțu
 - e) William Petty
10. Ca disciplină științifică, statistica are
- a) un pronunțat caracter metodologic
 - b) un pronunțat caracter teoretic
 - c) o sferă de cunoaștere limitată la nivelul individului
 - d) ca sferă de investigare fenomenele repetabile
11. Obiectul de studiu al statisticii îl constituie
- a) mișcările curente – continue ale fenomenelor și proceselor ce își au existența sub formă de colectivități
 - b) latura cantitativă a colectivităților statistice
 - c) latura calitativă a colectivităților statistice
 - d) aspectul stochastic al populației
12. Originalitatea metodei statistice constă în
- a) conjuncția variabilității fenomenelor cu rigoarea numerică a exprimării statistice
 - b) conjuncția variabilității fenomenelor cu caracterul dichotomic al unor variabile
 - c) estimarea parametrilor distribuțiilor empirice
 - d) estimarea parametrilor distribuțiilor teoretice

13. Legile cu caracter statistic se manifestă

- a) numai la nivelul întregului, ca urmare a acțiunii factorilor esențiali asupra variației
- b) numai la nivelul fiecărui caz în parte
- c) ca urmare a acțiunilor factorilor aleatori
- d) ca urmare a acțiunii factorilor aleatori și esențiali, la nivelul unei unități

14. Raționamentul statistic este

- a) numai un raționament inductiv
- b) numai un raționament deductiv
- c) un raționament deductiv-inductiv iterativ
- d) un raționament aleator

15. Principiile metodologice ale statisticii sunt

- a) observarea faptică
- b) exprimarea numerică
- c) autenticitatea datelor
- d) eficiența culegerii datelor

16. Particularitățile procesului cunoașterii statistice sunt date de

- a) folosirea unui raționament deductiv-inductiv, succesiv
- b) tratarea fenomenelor și proceselor ca pe un întreg structural
- c) tratarea fenomenelor numai de la general la particular
- d) tratarea fenomenelor numai de la particular la general

17. Cele mai noi tendințe în dezvoltarea statisticii vizează

- a) inferența neparametrică
- b) statistica robustă
- c) statistica descriptivă
- d) analiza datelor

18. Colectivitățile statistice sunt

- a) asociații finite de elemente unite între ele prin omogenitate
- b) asociații infinite de elemente unite între ele prin omogenitate
- c) elemente eterogene dintr-un spațiu dat
- d) grupuri de elemente diferențiate între ele după o caracteristică statistică

19. Clasificarea colectivităților statistice se poate face în funcție de

- a) natura elementelor componente
- b) gradul de variabilitate a caracteristicilor observate
- c) gradul de cuprindere a elementelor componente
- d) autenticitatea datelor înregistrate

20. Unitățile statistice sunt

- a) elementele componente ale unei colectivități
- b) elementele de observare și înregistrare a unei colectivități
- c) elementele care formează efectivul, volumul unei colectivități
- d) unități de măsură a variabilelor statistice

21. Variabilele statistice reprezintă

- a) trăsături esențiale purtate de toate elementele unei colectivități
- b) valoare înregistrată la nivelul unei unități statistice
- c) variantă de apariție a unei unități statistice
- d) probabilitatea de apariție a unei valori posibile x_i

22. Fenomenele și procesele din domeniul economico-social diferă între ele prin

- a) mărime sau formă
- b) frecvență sau intensitate
- c) înălțime și distanță
- d) coeziune sau interdependență

23. Din punctul de vedere al cuantificării, fenomenele pot fi

- a) direct măsurabile
- b) măsurabile cu aproximație
- c) măsurabile prin identificare

24. Elementele fizice pot fi

- a) direct măsurabile cantitativ
- b) măsurabile cu aproximație

25. Opinia unei persoane poate fi

- a) direct măsurabilă cantitativ
- b) măsurabilă cu aproximație

26. Operația de cuantificare presupune

- a) un set de reguli de atribuire a unei valori unităților statistice ale unei colectivități observate după o caracteristică
- b) un set de reguli de atribuire a unei valori caracteristicilor statistice ale unei colectivități
- c) un set de reguli de atribuire a unui simbol valorilor unei variabile statistice

27. Scala reprezintă

- a) instrumentul prin care se face diferențierea valorilor unei caracteristici

- b) un continuum de cifre sau de simboluri, plasate ierarhic, de la inferior la superior
- c) un instrument pentru calculul valorilor parametrilor unei distribuții

28. Tipurile de scală sunt

- a) nominală
- b) ordinală
- c) numerică (interval sau raport)

29. Scala nominală are ca proprietate

- a) identitatea categoriilor unei variabile
- b) ordinea categoriilor unei variabile
- c) diferențele între numere au un sens
- d) raportul dintre două puncte de scală este același

30. Proprietatea de identitate, specifică scalei nominale, presupune

- a) un interval între numere, care să permită compararea diferențelor dintre numere
- b) un raport de preferință, desemnat prin „>“, care permite măsurarea cu ranguri
- c) apartenența elementelor la o categorie

31. Scalarea cu ajutorul scalei nominale cere ca

- a) tuturor indivizilor unei categorii să li se atribuie aceeași valoare
- b) valorile numerice să fie atribuite unităților colectivității în funcție de sensul diferenței de mărime a caracteristicii observate
- c) doi indivizi care aparțin unor categorii distincte să aibă valori diferite
- d) între categoriile de unități ale colectivității să existe un raport de preferință

32. Codurile sunt

- a) identificatori, acordarea lor fiind pur convențională
- b) atribute acordate variabilelor cantitative
- c) valori numerice acordate în funcție de diferența de mărime a caracteristicii studiate

33. În cazul scalei nominale,

- a) între coduri există un raport de echivalență
- b) diferențele între două valori ale unei variabile nu au sens
- c) ordinea codurilor numerelor sau simbolurilor atribuite categoriilor este oarecare
- d) între categoriile variabilelor atributive există un raport de preferință

34. Scala ordinală sau cu ranguri are ca proprietate
- a) identitatea categoriilor unei variabile
 - b) ordinea categoriilor unei variabile
 - c) diferențele între numere au un sens
 - d) raportul dintre două puncte de scală este același
35. Proprietatea de ordine, specifică scalei ordinale, presupune
- a) ca între codurile acordate categoriilor unei caracteristici să fie un raport de echivalență
 - b) ca ordinea codurilor numerelor sau a simbolurilor atribuite categoriilor să fie oarecare
 - c) atribuirea de coduri numerice pentru fiecare categorie, clasă, atribut, în ordine descrescătoare, după importanța lor
36. Scala ordinală este folosită în cazul variabilelor atributive, când
- a) diferențele între două valori ale unei variabile au sens
 - b) între categorii există un raport de preferință, desemnat prin „>“, care permite măsurarea prin ranguri
 - c) diferențele între două valori ale unei variabile ordinale nu au sens
 - d) tuturor variabilelor li se atribuie aceeași valoare
37. Scala interval are ca proprietăți
- a) identitatea
 - b) ordinea
 - c) intervalul între numere are un sens
 - d) raportul dintre două puncte de pe scală este același
38. Scala interval este folosită în cazul
- a) variabilelor atributive
 - b) variabilelor nenumerice
 - c) variabilelor cantitative
39. Scala raport este folosită pentru cuantificarea
- a) variabilelor atributive
 - b) variabilelor nenumerice
 - c) variabilelor cantitative
40. Scala raport are ca proprietăți
- a) identitatea
 - b) ordinea
 - c) diferențele dintre numere au un sens
 - d) posedă un zero absolut, origine reală, fixă, ca punct de referință

41. Scala de intensitate este folosită pentru
- cuantificarea opiniilor, comportamentelor
 - variabilele cantitative
 - cuantificarea preferințelor indivizilor
42. Scala de opinie este o
- scală de intensitate
 - scală de rating
 - scală „cvasimetrică“
43. Scala de rating este
- o scală de opinie în trepte (4 până la 10 gradații)
 - scală cu 5 trepte: coincide perfect, coincide parțial, coincide într-o oarecare măsură, coincide mai puțin, nu coincide deloc
 - este o scală pentru cuantificarea variabilelor cantitative
44. Un indicator statistic este
- un rezultat al unei numărări
 - un rezultat numeric al unei măsurători statistice
 - un rezultat numeric al unui calcul asupra datelor înregistrate
45. Într-o cercetare statistică, un indicator statistic apare în ipostază de
- purtător de informații (reflectând, în expresie numerică, un fenomen real)
 - mijloc de calcul (un instrument metodologic de obținere a informațiilor statistice)
 - mijloc de codificare a variabilelor statistice
46. Indicatorii statistici trebuie să îndeplinească următoarele condiții de valabilitate științifică:
- să aibă un conținut științific bine determinat, o definiție sau o formulă de calcul
 - să fie comparabili în timp, în spațiu și din punct de vedere calitativ
 - să fie direct măsurabili și să coincidă, într-o oarecare măsură, cu fenomenul reflectat
47. Clasificarea indicatorilor se face după
- modul de determinare
 - gradul de cuprindere
 - forma de exprimare
48. După modul de determinare, se disting
- indicatori sintetici și analitici
 - indicatori statistici primari (elementari)

- c) indicatori statistici derivați

49. Indicatorii statistici primari reprezintă

- a) rezultatul unei prelucrări a indicatorilor primari
- b) expresii numerice ale categoriilor economice de sinteză
- c) rezultatul unei numărări, măsurători statistice și exprimă o dimensiune a unui element sau a unei colectivități

50. Indicatorii statistici derivați reprezintă

- a) rezultatul unei numărări, măsurători statistice și exprimă o dimensiune a unui element sau a unei colectivități
- b) rezultatul unei prelucrări a indicatorilor primari
- c) rezultatul unei codificări statistice

51. După gradul de cuprindere, indicatorii statistici sunt

- a) indicatori statistici sintetici
- b) indicatori primari
- c) indicatori statistici analitici

52. Indicatorii statistici sintetici reprezintă

- a) indicatori care exprimă componența pe grupe
- b) expresii numerice ale categoriilor economice de sinteză
- c) indicatori de rezultate la nivel macroeconomic (PIB, PN)

53. Indicatorii statistici analitici sunt

- a) indicatori care exprimă componența pe grupe sau structura colectivității
- b) expresii numerice ale categoriilor economice de sinteză
- c) indicatori care exprimă influența factorilor

54. După forma de exprimare, indicatorii statistici pot fi

- a) indicatori în mărimi absolute
- b) indicatori în mărimi relative
- c) indicatori în mărimi medii

55. Sistemul de indicatori statistici

- a) caracterizează multilateral fenomenele studiate de statistică
- b) măsoară și exprimă un aspect al fenomenelor studiate
- c) se agregă într-un sistem unic și complex de indicator



1.5. Recunoașteți exemplul corect

1. O unitate statistică poate fi
 - a) o persoană interogată într-o anchetă
 - b) o țară pentru care se studiază date socio-economice
 - c) salariul unei persoane angajate într-o firmă
 - d) ansamblul salariaților unei firme din localitatea A la un moment dat
2. O colectivitate statistică poate fi
 - a) fiecare persoană interogată într-o anchetă
 - b) ansamblul persoanelor interogate într-o anchetă
 - c) ansamblul salariaților unei firme din localitatea A la un moment dat
 - d) ansamblul țărilor pentru care se studiază date socio-economice
3. O colectivitate de stare poate fi
 - a) ansamblul persoanelor înregistrate pe teritoriul României la 18 martie 2002
 - b) opiniile electoratului înregistrate în noiembrie 2004
 - c) ansamblul cererilor de angajare depuse la firma B, în anul 2004
4. O colectivitate de mișcare poate fi
 - a) ansamblul persoanelor înregistrate pe teritoriul României la 18 martie 2002
 - b) opiniile electoratului înregistrate în noiembrie 2004
 - c) ansamblul cererilor de angajare depuse la firma B, în anul 2004
5. Sunt variabile calitative atributive (nominale)
 - a) talia, greutatea, salariul
 - b) sexul persoanelor, starea civilă, domiciliul
 - c) PNB/locuitor, speranța de viață, număr de locuitori ai unei țări
6. Sunt variabile calitative exprimate numeric
 - a) talia, greutatea, salariul
 - b) sexul persoanelor, starea civilă, domiciliul
 - c) PNB/locuitor, speranța de viață, număr de locuitori ai unei țări

7. Sunt variabile discrete

- a) PNB/locuitor, speranța de viață
- b) numărul de copii pe familie, numărul de camere ale unui apartament, numărul de piese defecte dintr-un lot de rulmenți
- c) sexul persoanelor, starea civilă, domiciliul

8. Sunt variabile continue

- a) talia, greutatea, salariul
- b) numărul de copii pe familie, numărul de camere ale unui apartament, numărul de piese defecte dintr-un lot de rulmenți
- c) ansamblul persoanelor interogate într-o anchetă

9. Atunci când se dispune de valorile luate de una sau mai multe variabile statistice, cunoașterea unei colectivități statistice presupune aplicarea

- a) statisticii descriptive
- b) statisticii inferențiale
- c) statisticii momentelor

10. Atunci când se dispune de valorile luate de una sau mai multe variabile statistice asupra unui eșantion, cunoașterea întregii colectivități din care este extras eșantionul, presupune folosirea

- a) statisticii descriptive
- b) statisticii inferențiale
- c) statisticii momentelor



1.5. Răspunsuri corecte

1.4. Alegeți răspunsul corect

1. a) Crearea cuvântului „statistică“ este atribuită lui Gottfried Achenwall.
2. b) De peste șase milenii, statistica este atestată ca activitate de strângere a datelor cu caracter militar și fiscal.
3. a) și b) În etapa „prestatistică“, statistica este cunoscută sub formă de numărătoare a populației și evidență a bunurilor materiale.
4. a) Instituționalizarea statisticii, în țara noastră, a avut loc în 1859, sub conducerea domnitorului Al.I. Cuza.
5. c) Ioan Ionescu de la Brad a scris *Povățuiri pentru facerea catagrafiei Moldovei, precedate de oarecare elemente de statistică* (1859).
6. b) Alexandru Bărbat a scris *Teoria statisticii sociale* (1972).
7. d) William Petty a scris *Political Arithmetic* (1676).
8. a), b), d) și e) Prin activitatea practică și opera lor, sunt considerați întemeietori ai statisticii în țara noastră Ioan Ionescu de la Brad, Dionisie Pop Marțian, Nicolae Șuțu și George Barițiu.
9. a), b), c) și e) Sunt considerați reprezentanți ai *Aritmeticii politice* Ioan Ionescu de la Brad, Edmund Halley, Dionisie Pop Marțian și William Petty.
10. a) Ca disciplină științifică, statistica are un pronunțat caracter metodologic.
11. a) Obiectul de studiu al statisticii îl constituie mișcările curente – continue ale fenomenelor și proceselor care își au existența sub formă de colectivități.
12. a) Originalitatea metodei statistice constă în conjuncția variabilității fenomenelor cu rigoarea numerică a exprimării statistice.
13. a) Legile cu caracter statistic se manifestă numai la nivelul întregului, ca urmare a acțiunii factorilor esențiali asupra variației.
14. c) Raționamentul statistic este un raționament deductiv-inductiv iterativ.

15. a) și b) Principiile metodologice ale statisticii sunt observarea faptică și exprimarea numerică.
16. a) și b) Particularitățile procesului cunoașterii statistice sunt date de folosirea unui raționament deductiv-inductiv, su cesiv și tratarea fenomenelor și proceselor ca pe un întreg structural.
17. a), b) și d) Cele mai noi tendințe în dezvoltarea statisticii vizează inferența neparametrică, statistica robustă și analiza datelor.
18. a) Colectivitățile statistice sunt asociații finite de elemente unite între ele prin omogenitate.
19. a) și c) Clasificarea colectivităților statistice se poate face în funcție de natura elementelor componente și gradul de cuprindere a elementelor componente.
20. a), b) și c) Unitățile statistice sunt elementele componente ale unei colectivități, elementele de observare și înregistrare a unei colectivități, precum și elementele care formează efectivul, volumul unei colectivități.
21. a) Variabilele statistice reprezintă trăsături esențiale purtate de toate elementele unei colectivități.
22. a), b) și d) Fenomenele și procesele din domeniul economico-social diferă între ele prin mărime sau formă, frecvență sau intensitate și coeziune sau interdependență.
23. a) și b) Din punctul de vedere al cuantificării, fenomenele pot fi direct măsurabile și măsurabile cu aproximație.
24. a) Elementele fizice pot fi direct măsurabile cantitativ.
25. b) Opinia unei persoane poate fi măsurabilă cu aproximație.
26. a) Operația de cuantificare presupune un set de reguli de atribuire a unei valori unităților statistice ale unei colectivități observate după o caracteristică.
27. a) și b) Scala reprezintă instrumentul prin care se face diferențierea valorilor unei caracteristici. Este un continuum de cifre sau de simboluri, plasate ierarhic, de la inferior la superior.
28. a), b) și c) Scalele pot fi nominale, ordinale, numerice (interval sau raport).
29. a) Scala nominală are ca proprietate identitatea categoriilor unei variabile.
30. c) Proprietatea de identitate, specifică scalei nominale, presupune apartenența elementelor la o categorie.

31. a) și c) Scalarea cu ajutorul scalei nominale cere ca tuturor indivizilor unei categorii să li se atribuie aceeași valoare, iar doi indivizi care aparțin unor categorii distincte să aibă valori diferite.
32. a) Codurile sunt identificatori, acordarea lor fiind pur convențională.
33. a) și c) În cazul scalei nominale, între coduri există un raport de echivalență, iar ordinea codurilor numerelor sau simbolurilor atribuite categoriilor este oarecare.
34. a) și b) Scala ordinală sau cu ranguri are ca proprietate identitatea categoriilor unei variabile și ordinea categoriilor unei variabile.
35. c) Proprietatea de ordine, specifică scalei ordinale, presupune ca atribuirea de coduri numerice pentru fiecare categorie, clasă, atribut, să se realizeze în ordine descrescătoare, după importanța lor.
36. b) și c) Scala ordinală este folosită în cazul variabilelor atributive, când între categorii există un raport de preferință, desemnat prin „>”, care permite măsurarea prin ranguri. Diferențele între două valori ale unei variabile ordinale nu au sens.
37. a), b) și c) Scala interval are ca proprietăți identitatea, ordinea și faptul că intervalul între numere are un sens.
38. c) Scala interval este folosită în cazul variabilelor cantitative (numerice).
39. c) Scala raport este folosită pentru cuantificarea variabilelor cantitative.
40. a), b), c) și d) Scala raport are ca proprietăți identitatea, ordinea, diferențele dintre numere au un sens, iar aceasta (scala raport) posedă un zero absolut, origine reală, fixă, ca punct de referință.
41. a) Scala de intensitate este folosită pentru cuantificarea opiniilor, comportamentelor.
42. a), b) și c) Scala de opinie este o scală de intensitate, de rating sau „cvasimetrică”.
43. a) și b) Scala de rating este o scală de opinie în trepte (4 până la 10 gradații). De exemplu, pentru o scală cu 5 trepte gradațiile pot fi formulate astfel: coincide perfect, coincide parțial, coincide într-o oarecare măsură, coincide mai puțin, nu coincide deloc.
44. a), b) și c) Un indicator statistic este un rezultat al unei numărări, un rezultat numeric al unei măsurători statistice și un rezultat numeric al unui calcul asupra datelor înregistrate.

45. a) și b) Într-o cercetare statistică, un indicator statistic apare în ipostaza de purtător de informații (reflectând în expresie numerică un fenomen real) și de mijloc de calcul (un instrument metodologic de obținere a informațiilor statistice).
46. a) și b) Indicatorii statistici trebuie să îndeplinească următoarele condiții de valabilitate științifică: să aibă un conținut științific bine determinat, o definiție sau o formulă de calcul și să fie comparabili în timp, în spațiu și din punct de vedere calitativ.
47. a), b) și c) Clasificarea indicatorilor se face după modul de determinare, gradul de cuprindere și forma de exprimare.
48. b) și c) După modul de determinare, se disting indicatori statistici primari (elementari) și indicatori statistici derivați.
49. c) Indicatorii statistici primari reprezintă rezultatul unei numărări, al unei măsurători statistice și exprimă o dimensiune a unui element sau a unei colectivități.
50. b) Indicatorii statistici derivați reprezintă rezultatul unei prelucrări a indicatorilor primari.
51. a) și c) După gradul de cuprindere, indicatorii statistici sunt indicatori statistici sintetici și indicatori statistici analitici.
52. b) și c) Indicatorii statistici sintetici reprezintă expresii numerice ale categoriilor economice de sinteză și indicatori de rezultate la nivel macroeconomic (PIB, PN).
53. a) și c) Indicatorii statistici analitici sunt indicatori care exprimă componența pe grupe sau structura colectivității și indicatori care exprimă influența factorilor.
54. a), b) și c) După forma de exprimare, indicatorii statistici pot fi indicatori în mărimi absolute, indicatori în mărimi relative și indicatori în mărimi medii.
55. a) și c) Sistemul de indicatori statistici caracterizează multilateral fenomenele studiate de statistică și se agregă într-un sistem unic și complex de indicatori.
56. c) Proprietatea de identitate, specifică relațiilor de agregare, este cea care asigură apartenența elementelor la o categorie.

1.5. Recunoașteți exemplul corect

1. a) și b) O unitate statistică poate fi o persoană interogată într-o anchetă și o țară pentru care se studiază date socio-economice.
2. b), c) și d) Colectivități statistice pot fi ansamblul persoanelor interogate într-o anchetă, ansamblul salariaților unei firme din localitatea „A” la un moment dat și ansamblul țărilor pentru care se studiază date socio-economice.
3. a) O colectivitate de stări poate fi ansamblul persoanelor înregistrate pe teritoriul României la 18 martie 2002.
4. b) și c) O colectivitate de mișcare poate fi ansamblul opiniilor electoratului înregistrate în noiembrie 2004, ansamblul cererilor de angajare depuse la firma B în anul 2004.
5. b) Variabile calitative atributive (nominale) sunt: sexul persoanelor, starea civilă, domiciliul.
6. a) și c) Variabile calitative exprimate numeric sunt: talia, greutatea, salariul, PNB/locuitor, speranța de viață, numărul de locuitori ai unei țări.
7. b) Variabile discrete sunt: numărul de copii pe o familie, numărul de camere ale unui apartament, numărul de piese defecte dintr-un lot de rulmenți.
8. a) Variabile continue sunt: talia, greutatea, salariul.
9. a) Cunoașterea unei colectivități statistice, atunci când se dispune de valorile luate de una sau mai multe variabile statistice asupra ansamblului, presupune aplicarea statisticii descriptive.
10. b) Atunci când se dispune de valorile luate de una sau mai multe variabile statistice asupra unui eșantion, cunoașterea întregii colectivități din care este extras eșantionul presupune folosirea statisticii inferențiale.

Capitolul 2

OBSERVAREA STATISTICĂ

- ◆ Principii de bază ale observării statistice
- ◆ Surse de date statistice
- ◆ Programul unei observări statistice
- ◆ Metode de înregistrare statistică
- ◆ Lucrări de înregistrare statistică
- ◆ Erori de înregistrare statistică
- ◆ Alegeți răspunsul corect
- ◆ Recunoașteți exemplul corect
- ◆ Răspunsuri corecte

2.1. Principii de bază ale observării statistice

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economică, București, 2002, pp. 29-30)

- Asigurarea autenticității datelor
- Eficiența observării

2.2. Surse de date statistice

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economică, București, 2002, pp. 30-32)

- **Sistemul informațional statistic (S.I.S)**
 - *Forme de organizare a S.I.S.*
 - *Sistemul informațional curent*
 - *Observări statistice special organizate*
 - *Publicații statistice*
 - *Bănci de date statistice*
- **Forme de organizare a S.I.S.**
 - *Sistem centralizat*
 - *Sistem descentralizat*

2.3. Programul unei observări statistice

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economică, București, 2002, pp. 32-35)

- **Probleme metodologice ale programului unei observări statistice**
 - *Delimitarea colectivității*
 - *Definirea unităților statistice*
 - *Alegerea variabilelor (caracteristicilor) ce urmează a fi înregistrate*
 - *Elaborarea formularelor statistice*
 - *Alegerea timpului și locului de înregistrare*

- **Probleme organizatorice**
 - *Recrutarea și instruirea anchetatorilor de teren*
 - *Organizarea teritoriului în care este delimitată colectivitatea observată*
 - *Organizarea activităților legate de procesul de înregistrare*
 - *Elaborarea devizului lucrării și asigurarea fondurilor necesare*

2.4. Metode de înregistrare statistică

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 35-37)

- **Înregistrări totale și înregistrări parțiale**
- **Înregistrări curente, periodice și ocazionale**

2.5. Lucrări de înregistrare statistică

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 37-43)

- **Recensământul** reprezintă o lucrare de înregistrare exhaustivă, la un moment dat, a unei colectivități de stări.
- **Rapoartele statistice**
 - *Rapoartele statistice reprezintă lucrări de înregistrare totală bazate pe criteriul cronologic al apariției și manifestării unităților din colectivitate.*
- **Anchetele prin sondaj**
 - *Anchetele prin sondaj presupun înregistrări parțiale, adică înregistrarea unui eșantion reprezentativ extras după principiile selecției din colectivitate totală.*

- **Monografia statistică**

- Monografia statistică este o înregistrare special organizată și are ca obiectiv cunoașterea multilaterală și în profunzime a unei singure unități complexe sau a unei singure probleme.

2.6. Erori de înregistrare statistică

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 43-48)

- **Noțiunea de eroare de înregistrare**

- Eroarea de înregistrare reprezintă diferența dintre rezultatul obținut prin înregistrare și mărimea concretă, reală a nivelurilor caracteristicilor înregistrate.

- **Tipuri de erori de înregistrare**

- Erori întâmplătoare, erori sistematice, greșeli
- Erori de acoperire și erori de reprezentativitate

- **Surse de erori de înregistrare**

- Unitatea de observare
- Limitele puterii de observare umană
- Definierea vagă a unității de observare sau a caracteristicilor de înregistrare
- Imperfecțiuni ale metodelor și mijloacelor de observare

- **Prevenirea și controlul statistic al erorilor de înregistrare**

- Prevenirea erorilor
- Controlul statistic



2.7. Alegeți răspunsul corect

1. Principiile de bază ale observării statistice sunt
 - a) asigurarea autenticității datelor
 - b) însumarea caracteristicilor de grupare
 - c) eficiența procesului de observare
2. Asigurarea autenticității datelor presupune
 - a) determinarea tendinței centrale a distribuției
 - b) înregistrarea datelor în ordine crescătoare
 - c) asigurarea concordanței dintre datele înregistrate și dimensiunea reală a fenomenelor
3. Condițiile cerute de principiul autenticității datelor sunt
 - a) condiția de volum
 - b) condiția de calitate
 - c) obținerea informațiilor în timp util
4. Eficiența procesului de observare presupune
 - a) să se culegă numai datele necesare, în concordanță cu programul observării
 - b) să se înregistreze datele în mod selectiv
 - c) să se determine indicatorii variației
5. Sursele de date statistice sunt
 - a) sistemul informațional curent
 - b) observările special organizate
 - c) publicațiile statistice
 - d) băncile de date statistice
6. Sistemul informațional statistic (S.I.S.) se compune din
 - a) ansamblul datelor statistice înscrise și păstrate pe suporturi de tip document și/sau magnetice
 - b) structuri clasice
 - c) informația de descriere a datelor statistice și a fluxurilor informaționale

7. Sistemul metainformațional statistic reprezintă

- a) informația de descriere a datelor statistice și a fluxurilor informaționale
- b) ansamblul datelor statistice înscrise și păstrate pe suporturi de tip document și/sau magnetice
- c) structuri clasice

8. Sistemul metainformațional statistic cuprinde

- a) clasificări
- b) nomenclatoare
- c) norme metodologice statistice și norme tehnice
- d) documentația de prelucrare statistică

9. Sistemul informațional statistic curent asigură

- a) culegerea, în mod curent, continuu, prin rapoarte statistice, a datelor înscrise în evidența primară, tehnico-operativă, contabilă și statistică
- b) accesul la date, cu viteza necesară și în mod repetat, ori de câte ori este nevoie
- c) culegerea datelor (care nu se găsesc în materialul faptic vechi) care constituie materialul faptic nou

10. Băncile de date asigură

- a) accesul la date, cu viteza necesară și în mod repetat, ori de câte ori este nevoie
- b) memorarea datelor agregate periodic în baze de date statistice primare
- c) culegerea, în mod curent, continuu, prin rapoarte statistice, a datelor înscrise în evidența primară, tehnico-operativă, contabilă și statistică

11. Autonomia, confidențialitatea, transparența, relevanța, proporționalitatea, deontologia statistică și principiul raportului cost/ eficiență reprezintă

- a) principii ale sistematizării datelor
- b) principii pe care se bazează organizarea și funcționarea statisticii oficiale în țara noastră
- c) principii de determinare a indicatorilor statistici

12. Programul unei observări statistice cuprinde

- a) modul de calcul al indicatorilor dispersiei
- b) metodologia procesului de culegere a datelor
- c) modul de organizare a procesului de culegere a datelor

13. Momentele procesului observării statistice sunt
- identificarea elementelor
 - măsurarea caracteristicilor
 - înregistrarea fiecărei unități cu valoarea caracteristicii observate
14. Problemele metodologice ale programului unei observări statistice sunt
- delimitarea colectivităților statistice
 - definirea unităților statistice
 - alegerea caracteristicilor care urmează a fi înregistrate
 - elaborarea formularelor necesare
 - alegerea timpului și locului de înregistrare
15. Problemele organizatorice ale procesului observării statistice sunt
- recrutarea și instruirea anchetatorilor
 - organizarea teritoriului în care este delimitată colectivitatea
 - organizarea activităților legate de procesul înregistrării
 - elaborarea devizului lucrărilor necesare și asigurarea fondurilor necesare
16. Metodele de înregistrare se clasifică în funcție de
- gradul de cuprindere a colectivității
 - natura colectivității
 - timpul în care se înregistrează datele
17. După gradul de cuprindere, înregistrările pot fi
- înregistrări totale
 - înregistrări periodice
 - înregistrări parțiale
18. După natura și timpul înregistrării, se disting
- înregistrări curente
 - înregistrări periodice (de momente)
 - înregistrări ocazionale
19. Lucrări de înregistrare statistică totală sunt
- anchetele prin sondaj
 - recensământul
 - rapoartele statistice
20. Lucrări de înregistrare statistică parțiale sunt
- anchetele prin sondaj
 - recensământul
 - rapoartele statistice

21. Recensământul este o lucrare de înregistrare statistică
- a) exhaustivă
 - b) la un moment dat
 - c) a unei colectivități de stări
22. Principiile și caracteristicile tehnice de culegere a datelor prin recensământ sunt
- a) universalitatea
 - b) definirea unităților statistice
 - c) periodicitatea și simultaneitatea recensământului
 - d) comparabilitatea datelor
 - e) caracterul nelimitativ al caracteristicilor din program
 - f) caracterul științific și aplicativ general
23. Programul recensământului vizează
- a) date pentru identificarea precisă a persoanelor
 - b) date despre caracteristicile demografice, sociale, economice, culturale etc.
 - c) date necesare organelor guvernamentale
24. Primele recensăminte moderne s-au organizat în țara noastră
- a) în anii 1860-1861
 - b) în anii 1870-1871
 - c) în anii 1859-1860
25. Înregistrările prin sondaj pot fi folosite pentru
- a) studierea bugetelor de familie
 - b) studierea problemelor pieței
 - c) controlul calității produselor și proceselor tehnologice
 - d) studierea opiniei
26. Instituții de studiere a opiniei publice sunt
- a) GALLUP (SUA, 1935)
 - b) IFOP (Franța, 1938)
 - c) IRSOP (România, 1990)
27. Eroarea de înregistrare reprezintă
- a) diferența dintre valoarea unei statistici și valoarea parametrului considerat
 - b) diferența dintre doi parametri comparabili
 - c) diferența dintre rezultatul obținut prin înregistrare și mărimea reală a nivelurilor caracteristicilor de înregistrare, adică $e_i = x - x_a$,
 $i = \overline{1, n}$, x_i – valoare înregistrată, x_a – valoare adevărată

28. Mărimea erorilor de înregistrare este direct proporțională cu
- volumul înregistrărilor
 - precizia mijloacelor de înregistrare
 - timpul de înregistrare
29. Surse de erori sunt
- toate elementele implicate în procesul de înregistrare
 - starea socială a anchetatorilor
 - unitatea de observare, cel care face înregistrarea, metoda și instrumentul de observare
30. Erorile datorate unității observate pot fi explicate prin
- limitele puterii de observare umană
 - variația sau inconstanța în timp a unității de observat
 - imperfecțiuni ale formularului de înregistrare
31. Erorile datorate celui care face observarea pot fi explicate prin
- limitele puterii de observare umană
 - variația sau inconstanța în timp a unității de observat
 - imperfecțiuni ale instrumentelor de înregistrare
32. Erorile datorate metodei și mijloacelor de observare pot fi explicate prin
- imperfecțiuni ale formularului de înregistrare
 - imperfecțiuni ale instrumentelor de înregistrare
 - variația sau inconstanța în timp a unității de observat
33. După sursa și modul de producere, erorile de înregistrare sunt
- erori întâmplătoare
 - erori sistematice
 - greșeli
34. Erorile de înregistrare întâmplătoare sunt
- erori care diferă între ele prin mărime și prin semn
 - erori care urmează legi de repartiție cunoscute
 - erori care se produc într-un singur sens
35. Erorile sistematice sunt
- erori care se produc într-un singur sens și care, prin acțiunea lor unilaterală, influențează asupra valorii indicatorilor
 - erori datorate acțiunii factorilor determinanți
 - erori datorate acțiunii factorilor aleatori

36. Greșelile sunt

- a) erori care urmează legi de repartiție cunoscute
- b) erori datorate lipsei de experiență, incompetenței
- c) erori care nu pot fi evaluate și nici depistate

37. Erorile de acoperire sunt

- a) erori datorate neconcordanței bazei de sondaj cu domeniul de delimitare a colectivității studiate
- b) erori sistematice
- c) erori care urmează legi de repartiție cunoscute

38. Erorile de reprezentativitate sunt

- a) erori datorate neconcordanței bazei de sondaj cu domeniul de delimitare a colectivității studiate
- b) abateri între valorile statistice calculate la nivel de eșantion și parametrii colectivității din care a fost extras eșantionul
- c) erori datorate lipsei de experiență, incompetenței



2.3. Recunoașteți exemplul corect

1. Atunci când se înregistrează zilnic, pentru o întreprindere, producția obținută, se efectuează
 - a) un recensământ
 - b) o înregistrare curentă-continuuă
 - c) o anchetă de opinie
2. Atunci când se înregistrează la un moment critic populația unei țări se efectuează
 - a) un recensământ
 - b) o înregistrare curentă-continuuă
 - c) o anchetă de opinie
3. Înregistrarea opiniei electoratului se realizează prin
 - a) recensământ
 - b) înregistrare curentă-continuuă
 - c) sondaj statistic
4. Se observă populația unei țări la 10 momente de recensământ. Colectivitatea studiată este reprezentată de
 - a) momentele de recensământ
 - b) populația țării
 - c) țara
5. Se observă populația unei țări la 10 momente de recensământ. Variabila studiată este
 - a) momentele de recensământ
 - b) populația țării
 - c) țara



2.9. Răspunsuri corecte

2.7. Alegeți răspunsul corect

1. a) și c) Principiile de bază ale observării statistice sunt asigurarea autenticității datelor și eficiența procesului de observare.
2. c) Asigurarea autenticității datelor presupune asigurarea concordanței dintre datele înregistrate și dimensiunea reală a fenomenelor.
3. a), b) și c) Condițiile cerute de principiul autenticității datelor sunt condiția de volum, condiția de calitate și obținerea informațiilor în timp util.
4. a) Eficiența procesului de observare presupune să se culegă numai datele necesare în concordanță cu programul observării.
5. a), b), c) și d) Sursele de date statistice sunt sistemul informațional curent, observările special organizate, publicațiile statistice și băncile de date statistice.
6. a) și c) Sistemul informațional statistic (S.I.S.) se compune din ansamblul datelor statistice înscrise și păstrate pe suporturi de tip document și/sau magnetice și din informația de descriere a datelor statistice și a fluxurilor informaționale.
7. a) Sistemul metainformațional statistic reprezintă informația de descriere a datelor statistice și a fluxurilor informaționale.
8. a), b), c) și d) Sistemul metainformațional statistic cuprinde clasificări, nomenclatoare, norme metodologice statistice și norme tehnice, precum și documentația de prelucrare statistică.
9. a) Sistemul informațional statistic curent asigură culegerea, în mod curent, continuu, prin rapoarte statistice, a datelor înscrise în evidența primară, tehnico-operativă, contabilă și statistică.
10. a) și b) Băncile de date asigură accesul la date, cu viteza necesară și în mod repetat, ori de câte ori este nevoie și memorarea datelor agregate periodic în baze de date statistice primare.

11. b) Autonomia, confidențialitatea, transparența, relevanța, proporționalitatea, deontologia statistică și principiul raportului cost/eficiență reprezintă principii pe care se bazează organizarea și funcționarea statisticii oficiale în țara noastră.
12. b) și c) Programul unei observări statistice cuprinde metodologia procesului de culegere a datelor și modul de organizare a procesului de culegere a datelor.
13. a), b) și c) Momentele procesului observării statistice sunt identificarea elementelor, măsurarea caracteristicilor și înregistrarea fiecărei unități, cu valoarea caracteristicii observate.
14. a), b), c), d) și e) Problemele metodologice ale programului unei observări statistice sunt delimitarea colectivităților statistice, definirea unităților statistice, alegerea caracteristicilor care urmează a fi înregistrate, elaborarea formularelor necesare, alegerea timpului și locului de înregistrare.
15. a), b), c) și d) Problemele organizatorice ale procesului observării statistice sunt recrutarea și instruirea anchetatorilor, organizarea teritoriului în care este delimitată colectivitatea, organizarea activităților legate de procesul înregistrării, elaborarea devizului lucrărilor necesare și asigurarea fondurilor necesare.
16. a), b) și c) Metodele de înregistrare se clasifică în funcție de gradul de cuprindere a colectivității, natura colectivității și timpul în care se înregistrează datele.
17. a) și c) După gradul de cuprindere, înregistrările pot fi înregistrări totale și înregistrări parțiale.
18. a), b) și c) După natura și timpul înregistrării, se disting înregistrări curente, înregistrări periodice (de momente) și înregistrări ocazionale.
19. b) și c) Lucrări de înregistrare statistică totală sunt recensământul și rapoartele statistice.
20. a) Lucrări de înregistrare statistică parțială sunt anchetele prin sondaj.
21. a), b) și c) Recensământul este o lucrare de înregistrare statistică exhaustivă a unei colectivități de stări, care se efectuează la un moment dat.
22. a), c), d), e) și f) Principiile și caracteristicile tehnice de culegere a datelor prin recensământ sunt: universalitatea, periodicitatea și simultaneitatea recensământului, comparabilitatea datelor, caracterul nelimitativ al caracteristicilor din program, caracterul științific și aplicativ general.

23. a), b) și c) Programul recensământului vizează date pentru identificarea precisă a persoanelor, date despre caracteristicile demografice, sociale, economice, culturale etc. și date necesare organelor guvernamentale.
24. c) Primele recensăminte moderne s-au organizat în țara noastră în anii 1859-1860.
25. a), b), c) și d) Înregistrările prin sondaj pot fi folosite pentru studierea bugetelor de familie, studierea problemelor pieței, controlul calității produselor și proceselor tehnologice și studierea opiniei.
26. a), b) și c) Instituții de studiere a opiniei publice sunt GALLUP (SUA, 1935), IFOP (Franța, 1938) și IRSOP (România, 1990).
27. c) Eroarea de înregistrare reprezintă diferența dintre rezultatul obținut prin înregistrare și mărimea reală a nivelurilor caracteristicilor de înregistrare, adică $e_i = x - x_a$, $i = \overline{1, n}$, x_i – valoare înregistrată, x_a – valoare adevărată.
28. a) și b) Mărimea erorilor de înregistrare este direct proporțională cu volumul înregistrărilor și precizia mijloacelor de înregistrare.
29. a) și c) Surse de erori sunt toate elementele implicate în procesul de înregistrare, unitatea de observare, cel care face înregistrarea, metoda și instrumentul de observare.
30. b) Erorile datorate unității observate pot fi explicate prin variația sau inconstanța în timp a unității de observat.
31. a) Erorile datorate celui care face observarea pot fi explicate prin limitele puterii de observare umană.
32. a) și b) Erorile datorate metodei și mijloacelor de observare pot fi explicate prin imperfecțiuni ale formularului de înregistrare și imperfecțiuni ale instrumentelor de înregistrare.
33. a), b) și c) După sursa și modul de producere, erorile de înregistrare sunt erori întâmplătoare, sistematice și greșeli.
34. a) și b) Erorile de înregistrare întâmplătoare sunt erori care diferă între ele prin mărime și prin semn și erori care urmează legi de repartiție cunoscute.
35. a) și b) Erorile sistematice sunt erori care se produc într-un singur sens și care, prin acțiunea lor unilaterală, influențează asupra valorii indicatorilor și erori datorate acțiunii factorilor determinanți.
36. b) și c) Greșelile sunt erori datorate lipsei de experiență, incompetenței și erori care nu pot fi evaluate și nici depistate.

37. a) Erorile de acoperire sunt erori datorate neconcordanței bazei de sondaj cu domeniul de delimitare a colectivității studiate.
38. b) Erorile de reprezentativitate sunt abateri între valorile statistice calculate la nivel de eșantion și parametrii colectivității din care a fost extras eșantionul.

2.8. Recunoașteți exemplul corect

1. b) Atunci când se înregistrează zilnic, pentru o întreprindere, producția obținută, se efectuează o înregistrare curentă-continuu.
2. a) Atunci când se înregistrează la un moment critic populația unei țări, se efectuează un recensământ.
3. c) Înregistrarea opiniei electoratului se realizează prin sondaj statistic.
4. b) Se observă populația unei țări la 10 momente de recensământ. Colectivitatea studiată este populația țării.
5. a) Se observă populația unei țări la 10 momente de recensământ. Variabila studiată este reprezentată de momentele de recensământ.

27. a) Erorile de neapreciere sunt erorile datorate neaprecierii corecte a valorilor
 prezente în populație și care pot fi evitate prin utilizarea metodei de selecție aleatorie.

28. b) Erorile de neapreciere sunt erorile datorate neaprecierii corecte a valorilor
 prezente în populație și care pot fi evitate prin utilizarea metodei de selecție aleatorie.

25. a), b), c) și d) Înregistrările pot fi înregistrate în mod corect prin utilizarea
 bugetelor de familie, bugetelor de familie, bugetelor de familie, bugetelor de familie
 produselor și proceselor tehnologice.

26. a) și b) Înregistrările pot fi înregistrate în mod corect prin utilizarea
 bugetelor de familie, bugetelor de familie, bugetelor de familie, bugetelor de familie.

27. c) Erorile de neapreciere sunt erorile datorate neaprecierii corecte a valorilor
 prezente în populație și care pot fi evitate prin utilizarea metodei de selecție aleatorie.

28. a) și b) Erorile de neapreciere sunt erorile datorate neaprecierii corecte a valorilor
 prezente în populație și care pot fi evitate prin utilizarea metodei de selecție aleatorie.

29. a) și b) Erorile de neapreciere sunt erorile datorate neaprecierii corecte a valorilor
 prezente în populație și care pot fi evitate prin utilizarea metodei de selecție aleatorie.

30. b) Erorile datorate unității de observare pot fi explicate prin variația sau
 inconstanța în timp a unității de observare.

31. a) Erorile datorate unității de observare pot fi explicate prin variația sau
 inconstanța în timp a unității de observare.

32. a) și b) Erorile datorate metodei de observare pot fi explicate prin variația sau
 inconstanța în timp a unității de observare.

33. a), b) și c) După surse și moduri de producere, erorile de înregistrare sunt
 erori întâmplătoare, sistematice și greșeli.

34. a) și b) Erorile de înregistrare întâmplătoare sunt erori care diferă între ele
 prin mărime și prin semn și care sunt urmări ale legii de repartiție cunoscută.

35. a) și b) Erorile sistematice sunt erori care se pot observa într-un anumit sens și
 care sunt urmări ale legii de repartiție cunoscută.

36. b) și c) Greșelile sunt erori datorate lipsei de experiență, incompetenței și
 erori care pot fi evitate prin utilizarea metodei de selecție aleatorie.

Capitolul 3

SISTEMATIZAREA ȘI PREZENTAREA DATELOR STATISTICE

- ◆ Sistematizarea datelor
- ◆ Prezentarea datelor statistice
- ◆ Prezentarea distribuțiilor statistice unidimensionale
- ◆ Prezentarea distribuțiilor statistice bidimensionale
- ◆ Reprezentarea grafică a unei distribuții în SPSS
- ◆ Alegeți răspunsul corect
- ◆ Recunoașteți exemplul corect
- ◆ Răspunsuri corecte
- ◆ Probleme

3.1. Sistematizarea datelor

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 50-57)

- **Procedee de sistematizare**
 - *Centralizarea statistică a datelor*
 - *Gruparea statistică a datelor*
- **Tipuri de grupări statistice**
 - *După numărul caracteristicilor de grupare:*
 - *gruparea simplă*
 - *gruparea combinată*
 - *După natura caracteristicii de grupare:*
 - *gruparea după o variabilă de timp*
 - *gruparea după o variabilă de spațiu*
 - *gruparea după o variabilă calitativă (numerică sau atributivă)*
- **Probleme ale grupării statistice**
 - *Precizarea scopului grupării statistice*
 - *Alegerea variabilei/variabilelor de grupare*
 - *Stabilirea numărului de grupe în care urmează să se împartă unitățile colectivității*
 - *Determinarea mărimii intervalelor de grupare, în cazul variabilelor exprimate numeric*
 - *Delimitarea grupelor de variație și separarea unităților colectivității pe intervale de grupare (pe grupe)*

3.2. Prezentarea datelor statistice

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 57-61)

- **Tabele statistice**
 - *Elementele unui tabel statistic*
 - *titlul general*
 - *titlurile interioare*

- unitatea de măsură utilizată
- notele explicative
- sursa datelor
- rubricile tabelului
- *Tipuri de tabele statistice*
 - tabelul simplu
 - tabelul cu dublă intrare (tabel de corelație, tabel de asociere)

- **Reprezentări grafice**

- *Elementele unei reprezentări grafice*
 - axele de coordonate
 - scara
 - rețeaua graficului
 - legenda
- *Tipuri de reprezentări grafice*
 - pentru serii de timp
 - pentru serii de spațiu
 - pentru serii calitative

3.3. Prezentarea distribuțiilor statistice unidimensionale

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 62-84)

- **Definiții. Notății**

- O distribuție statistică unidimensională prezintă *corespondența dintre două șiruri de date statistice, sistematizate într-o succesiune logică: primul șir reprezintă valori ale caracteristicii de grupare, iar al doilea șir reprezintă frecvența de apariție corespunzătoare.*

- O distribuție statistică cu frecvențe egale

$X : (x_1, x_2, \dots, x_i, \dots, x_m)$ sau $X : (x_i)$ cu $i = 1, m$,

când $n_1 = n_2 = \dots = n_i$, respectiv,

- O distribuție statistică cu frecvențe diferite

$X : \begin{pmatrix} x_1, x_2, \dots, x_i, \dots, x_m \\ n_1, n_2, \dots, n_i, \dots, n_m \end{pmatrix}$ sau $X : \begin{pmatrix} x_i \\ n_i \end{pmatrix}$, cu $i = 1, m$,

când $n_1 \neq n_2 \neq \dots \neq n_i$.

- O distribuție statistică de intervale

$X : (J_i, n_i)$, unde $i = 1, k$, unde J_i reprezintă intervalele (x_{i-1}, x_i) , $i=1, k$ clase sau intervale de valori.

- *Frecvențe relative și frecvențe cumulate*
- *Prezentarea tabelară a unei distribuții unidimensionale*

- **Reprezentarea grafică a distribuțiilor statistice unidimensionale**

- *Poligonul și curba frecvențelor*
- *Curba frecvențelor cumulate*
- *Histograma*
- *Diagrame de structură*
- *Cronograme și diagrame polare*
- *Cartograme și cartodiagrame*

3.4. Prezentarea distribuțiilor statistice bidimensionale

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 62-84)

- **Definiții. Notății**

- O distribuție statistică bidimensională prezintă *variația unităților unei colectivități simultan după două caracteristici de grupare.*

- O distribuție statistică bidimensională este definită de ansamblul de triplete:

(x_i, y_j, n_{ij}) , $i=1, m, j=1, p$.

- *Frecvențe relative marginale, parțiale și condiționate*
- *Prezentarea tabelară a unei distribuții bidimensionale*

- **Reprezentarea grafică**

- *Norul de puncte*
- *Procedeul paralelipipedelor și al suprafeței poliedrale*
- *Piramida vârstelor*
- *Diagrame de structură construite în același plan*

3.5. Reprezentarea grafică a unei distribuții în SPSS

(Vezi Elisabeta Jaba și Ana Grama, *Analiza statistică cu SPSS sub Windows*, Editura Polirom, Iași, 2004, pp. 96-116)

- **Histograma și curba frecvențelor**
 - Construirea histogramei și a curbei frecvențelor presupune selectarea opțiunii *Histogram* din meniul *Graphs*. Reprezentarea grafică a acestora este prezentată în figura 3.1.

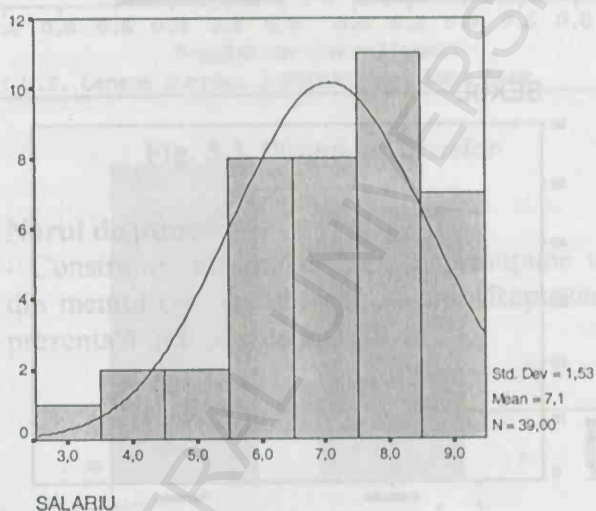


Fig. 3.1. Histograma

- **Diagrame de structură**
 - Construirea diagramelor de structură presupune următorul algoritm: din meniul *Analyse*, se alege *Descriptives Statistics*, opțiunea *Frequencies*, butonul de comandă *Charts*. Reprezentarea grafică a acestora este prezentată în figura 3.2.

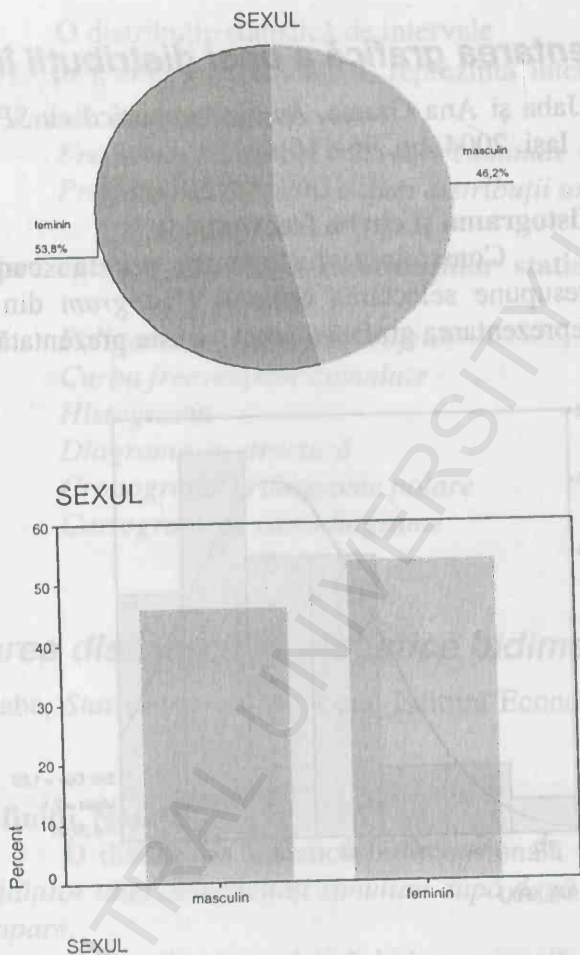


Fig. 3.2. Diagrame de structură

- **Piramida vârstelor**

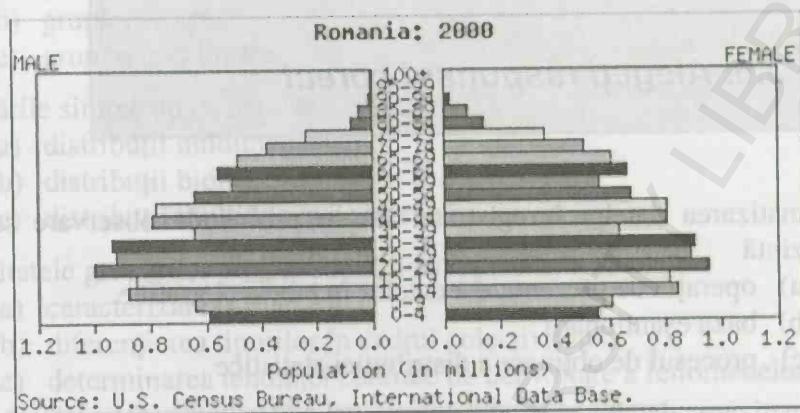


Fig. 3.3. Piramida vârstelor

- **Norul de puncte**

- Construirea norului de puncte presupune următorul demers: din meniul *Graphs* se alege *Scatter*. Reprezentarea grafică este prezentată în figura de mai jos.

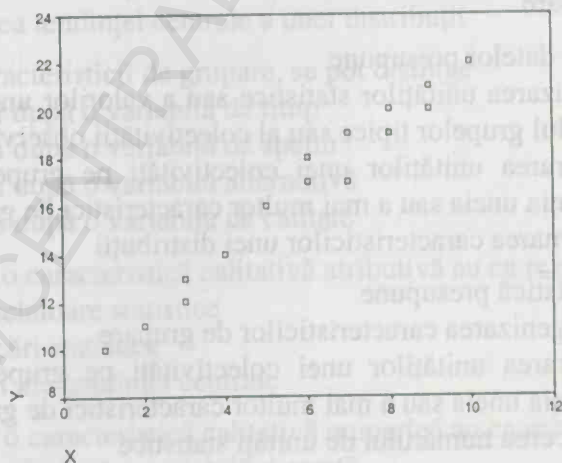


Fig. 3.4. Norul de puncte



3.6. Alegeți răspunsul corect

1. Sistematizarea datelor înregistrate într-un proces de observare statistică reprezintă
 - a) operația de prezentare a datelor în tabele și grafice
 - b) baza eșantionării
 - c) procesul de obținere a distribuției statistice
2. Sistematizarea datelor presupune un set de operații, și anume:
 - a) strângerea formularelor și desprinderea din formulare a datelor individuale
 - b) separarea unităților unei colectivități pe grupe eterogene
 - c) sortarea și totalizarea datelor
3. Sistematizarea datelor se poate realiza prin
 - a) estimare
 - b) centralizare
 - c) grupare
4. Centralizarea datelor presupune
 - a) totalizarea unităților statistice sau a valorilor unei caracteristici la nivelul grupelor tipice sau al colectivității observate
 - b) separarea unităților unei colectivități pe grupe omogene, după variația uneia sau a mai multor caracteristici de grupare
 - c) însumarea caracteristicilor unei distribuții
5. Gruparea statistică presupune
 - a) omogenizarea caracteristicilor de grupare
 - b) separarea unităților unei colectivități pe grupe omogene, după variația uneia sau a mai multor caracteristici de grupare
 - c) reducerea numărului de unități statistice
6. Gruparea datelor se poate efectua după
 - a) numărul caracteristicilor de grupare
 - b) volumul colectivității statistice
 - c) natura caracteristicilor de grupare
 - d) numărul de grupe

7. După numărul caracteristicilor de grupare se pot obține

- a) grupări sintetice
- b) grupări simple
- c) grupări combinate

8. Grupările simple au ca rezultat

- a) distribuții unidimensionale
- b) distribuții bidimensionale
- c) distribuții multidimensionale

9. Rezultatele grupărilor simple sunt folosite pentru

- a) caracterizarea structurii
- b) diferențierea tipurilor în cadrul colectivității
- c) determinarea tendinței centrale de dezvoltare a fenomenelor
- d) descrierea statistică a legăturilor dintre fenomene

10. Grupările combinate au ca rezultat

- a) distribuții unidimensionale
- b) distribuții bidimensionale
- c) distribuții multidimensionale

11. Grupările combinate sunt folosite pentru

- a) studierea legăturilor statistice
- b) studierea influenței factorilor asupra variației
- c) studierea tendinței centrale a unei distribuții

12. După natura caracteristicii de grupare, se pot distinge

- a) grupări după o variabilă de timp
- b) grupări după o variabilă de spațiu
- c) grupări după o variabilă alternativă
- d) grupări după o variabilă de calitate

13. Grupările după o caracteristică calitativă atributivă au ca rezultat

- a) nomenclatoare statistice
- b) clasificări statistice
- c) mărimi ale tendinței centrale

14. Grupările după o caracteristică calitativă numerică au ca rezultat

- a) distribuții după o variabilă discretă
- b) distribuții după o variabilă nominală
- c) distribuții după o variabilă continuă

15. Rezolvarea problemelor privind precizarea scopului grupării, alegerea variabilelor de grupare, stabilirea numărului de grupe, stabilirea mărimii intervalelor de grupare, delimitarea grupelor de variație și separarea unităților pe intervale de grupare reprezintă probleme ale
- a) grupării statistice
 - b) centralizării statistice
 - c) prezentării datelor în tabele statistice
16. Scopul grupării constă în
- a) analiza directă, în cazul grupărilor tipologice
 - b) sistematizarea materialului faptic în vederea prelucrării
 - c) alegerea caracteristicii de grupare
 - d) determinarea mărimii intervalelor de grupare
17. Alegerea variabilei/variabilelor de grupare se face în funcție de
- a) numărul de grupe în care se sistematizează datele
 - b) scopul grupării
 - c) esențialitatea caracteristicii de grupare
18. Stabilirea numărului de grupe și a mărimii intervalelor de grupare se realizează în funcție de
- a) scopul grupării
 - b) natura caracteristicilor de grupare
 - c) numărul unităților statistice
19. Numărul de grupe (k) se calculează după relația
- a) $k = \frac{l}{x_{\max} - x_{\min}}$
 - b) $k = A/l$
 - c) $k = 1 + 3.322 \log n$
20. Mărimea intervalelor de grupare se stabilește conform relației
- a) $l = A/k$
 - b) $l = \frac{x_{\min} - x_{\max}}{k}$
 - c) $l = A/(1 + 3.322 \log n)$
21. Intervalele de grupare pot fi
- a) deschise
 - b) închise
 - c) egale
 - d) inegale

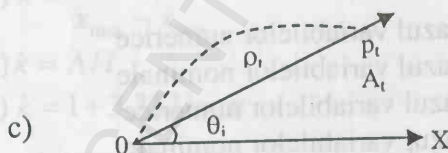
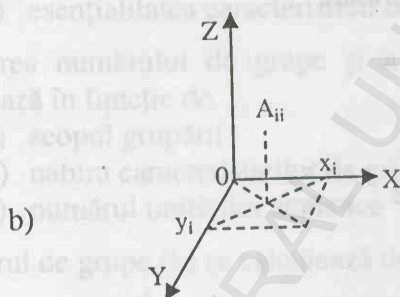
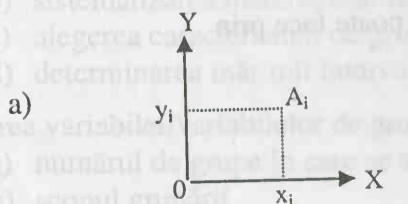
22. Gruparea pe variante de variație a caracteristicii/caracteristicilor de grupare presupune
- a) însumarea caracteristicilor de grupare
 - b) ordonarea numărului unităților statistice în sens crescător
 - c) ordonarea datelor în sens crescător sau descrescător
23. Gruparea pe intervale de variație presupune
- a) stabilirea mărimii intervalului de variație
 - b) delimitarea grupelor (intervalului) de variație
 - c) separarea unităților pe intervale de variație
24. Prezentarea rezultatelor grupării se poate face prin
- a) serii
 - b) tabele
 - c) grafice
25. Elementele unui tabel statistic sunt
- a) titlul general și titlurile interioare
 - b) rubricile
 - c) legenda
 - d) notele, sursa
26. Tabelele statistice pot fi
- a) tabele cu axe de coordonate
 - b) tabele simple
 - c) tabele cu dublă intrare
27. Tabelele cu dublă intrare sunt
- a) tabele de corelație, în cazul variabilelor numerice
 - b) tabele de corelație, în cazul variabilelor nominale
 - c) tabele de asociere, în cazul variabilelor numerice
 - d) tabele de asociere, în cazul variabilelor nominale
28. Reprezentarea grafică este
- a) o metodă de însumare a datelor
 - b) metodă de grupare a datelor
 - c) metodă de prezentare vizuală a datelor, într-un sistem de coordonate dat
29. Elementele unui grafic sunt
- a) titlul
 - b) axele de coordonate
 - c) scara, rețeaua

- d) diagrama
- e) legenda

30. Axele de coordonate pot fi

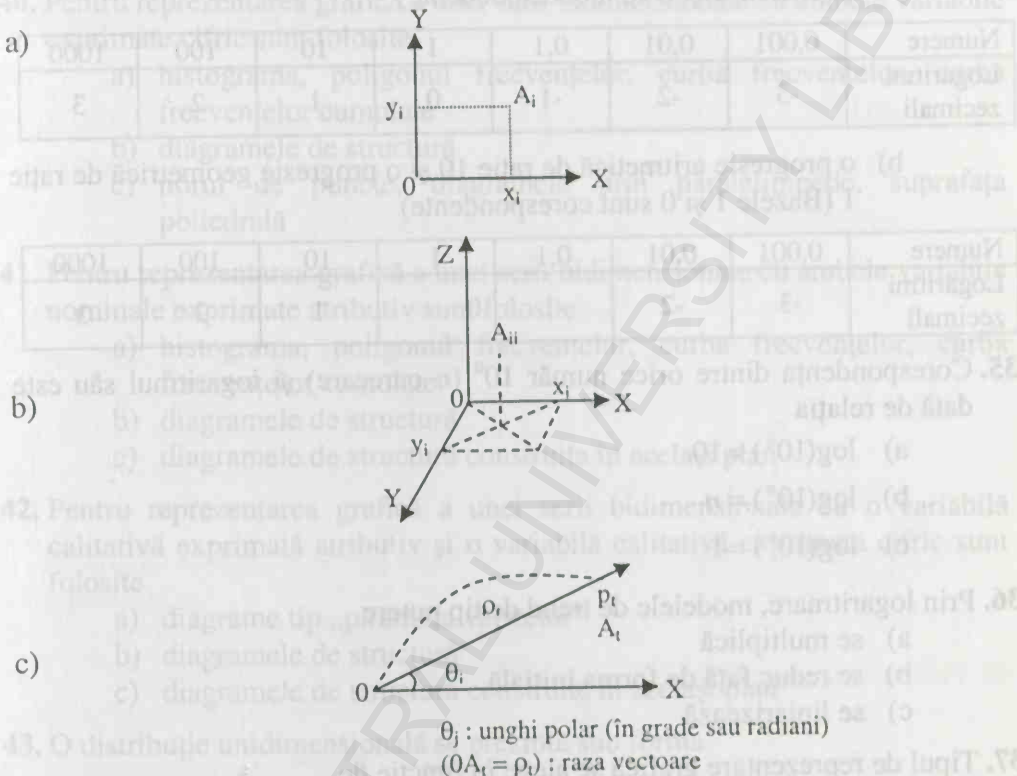
- a) axe în sistem de coordonate rectangulare
- b) axe în sistem de coordonate normale
- c) axe în sistem de coordonate polare

31. Reprezentarea unui punct într-un sistem de coordonate rectangulare este dată în figura



θ_j : unghi polar (în grade sau radiani)
 $(OA_t = p_t)$: raza vectorială

32. Reprezentarea unui punct într-un sistem de coordonate polare este dată în figura



33. Scara de reprezentare grafică poate fi

- a) uniformă (aritmetică)
- b) polară
- c) neuniformă (gaussiană)

34. Trecerea de la o scară la alta presupune realizarea unei corespondențe dintre

- a) o progresie geometrică de rație 10 și o progresie aritmetică de rație 1 (Bazele 1 și 0 sunt corespondente)

Numere	0,001	0,01	0,1	1	10	100	1000
Logaritmi zecimali	-3	-2	-1	0	1	2	3

- b) o progresie aritmetică de rație 10 și o progresie geometrică de rație 1 (Bazele 1 și 0 sunt corespondente)

Numere	0,001	0,01	0,1	1	10	100	1000
Logaritmi zecimali	-3	-2	-1	0	1	2	3

35. Corespondența dintre orice număr 10^n (n oarecare) și logaritmul său este dată de relația

- a) $\log(10^n) = 10$
 b) $\log(10^n) = n$
 c) $\log(10^n) = 1$

36. Prin logaritmare, modelele de trend de tip putere

- a) se multiplică
 b) se reduc față de forma inițială
 c) se liniarizează

37. Tipul de reprezentare grafică se alege în funcție de

- a) natura seriilor (distribuțiilor) statistice
 b) forma de exprimare a caracteristicilor de grupare
 c) numărul unităților statistice

38. Pentru reprezentarea grafică a unei serii calitative unidimensionale cu variabila exprimată numeric sunt folosite

- a) histograma, poligonul frecvențelor, curba frecvențelor, curba frecvențelor cumulate
 b) diagramele de structură
 c) norul de puncte, diagramele prin paralelipipede, suprafața poliedrală

39. Pentru reprezentarea grafică a unei serii calitative unidimensionale cu variabila exprimată atributiv sunt folosite

- a) histograma, poligonul frecvențelor, curba frecvențelor, curba frecvențelor cumulate

- b) diagramele de structură
- c) norul de puncte, diagramele prin paralelipiede, suprafața poliedrală

40. Pentru reprezentarea grafică a unei serii bidimensionale cu ambele variabile exprimate cifric sunt folosite

- a) histograma, poligonul frecvențelor, curba frecvențelor, curba frecvențelor cumulate
- b) diagramele de structură
- c) norul de puncte, diagramele prin paralelipiede, suprafața poliedrală

41. Pentru reprezentarea grafică a unei serii bidimensionale cu ambele variabile nominale exprimate atributiv sunt folosite

- a) histograma, poligonul frecvențelor, curba frecvențelor, curba frecvențelor cumulate
- b) diagramele de structură
- c) diagramele de structură construite în același plan

42. Pentru reprezentarea grafică a unei serii bidimensionale cu o variabilă calitativă exprimată atributiv și o variabilă calitativă exprimată cifric sunt folosite

- a) diagrame tip „piramida vârstelor“
- b) diagramele de structură
- c) diagramele de structură construite în același plan

43. O distribuție unidimensională se prezintă sub forma

- a) $X, Y : \{x_i, y_j, n_{ij}\}, i = \overline{1, m} \text{ și } j = \overline{1, p}$
- b) $X : \left\{ \begin{matrix} x_1 & x_2 & \dots & x_i & \dots & x_m \\ n_1 & n_2 & \dots & n_i & \dots & n_m \end{matrix} \right\}, i = \overline{1, m}$
- c) $X : (x_i), i = \overline{1, n}$

44. O distribuție bidimensională se prezintă sub forma

- a) $X, Y : \{x_i, y_j, n_{ij}\}, i = \overline{1, m} \text{ și } j = \overline{1, p}$
- b) $X : (x_i), i = \overline{1, n}$
- c) $X : \left\{ \begin{matrix} x_1 & x_2 & \dots & x_i & \dots & x_m \\ n_1 & n_2 & \dots & n_i & \dots & n_m \end{matrix} \right\}, i = \overline{1, m}$

45. O distribuție marginală se prezintă sub forma

- a) $\{x_i, n_{i\cdot}\}, i = \overline{1, m}$

$$b) \{x_{i,n_{\bullet j}}\}, j = \overline{1, p}$$

$$c) \{y_{j,n_{\bullet j}}\}, j = \overline{1, p}$$

$$d) \{y_{j,n_{i\bullet}}\}, i = \overline{1, m}$$

46. Distribuțiile condiționate se prezintă sub forma

$$a) (X|Y = y_j): \{x_{i,n_{ij}}\}, j = \overline{1, p}, i - \text{valoare fixa}$$

$$b) (X|Y = y_j): \{x_{i,n_{ij}}\}, i = \overline{1, m}, j - \text{valoare fixa}$$

$$c) (Y|X = x_i): \{y_{j,n_{ij}}\}, j = \overline{1, p}, i - \text{valoare fixa}$$

$$d) (Y|X = x_i): \{y_{j,n_{ij}}\}, i = \overline{1, m}, j - \text{valoare fixa}$$

47. Frecvențele relative marginale sunt definite de

$$a) f_{i\bullet} = \frac{n_{i\bullet}}{n_{\bullet\bullet}}$$

$$b) f_{ij} = \frac{n_{ij}}{n_{\bullet\bullet}}$$

$$c) f_{\bullet j} = \frac{n_{\bullet j}}{n_{\bullet\bullet}}$$

48. Frecvențele relative parțiale sunt definite de

$$a) f_{ij} = \frac{n_{ij}}{n_{\bullet\bullet}}$$

$$b) f_{i\bullet} = \frac{n_{i\bullet}}{n_{\bullet\bullet}}$$

$$c) f_{\bullet j} = \frac{n_{\bullet j}}{n_{\bullet\bullet}}$$

49. Frecvențele relative condiționate sunt definite de

$$a) f_{ij|j} = \frac{n_{ij}}{n_{\bullet j}}, j - \text{valoare fixa}, i = \overline{1, m}$$

$$b) f_{j|i} = \frac{n_{ij}}{n_{i\bullet}}, i - \text{valoare fixa}, j = \overline{1, p}$$

$$c) f_{ij} = \frac{n_{ij}}{n_{\bullet\bullet}}, i = \overline{1, m} \text{ și } j = \overline{1, p}$$



3.7. Recunoașteți exemplul corect

1. Distribuția populației ocupate a unei țări pe sectoare de activitate, la un moment dat, reprezintă o
 - a) distribuție de timp
 - b) distribuție de spațiu
 - c) distribuție calitativă cu variabila exprimată numeric
 - d) distribuție calitativă cu variabila exprimată atributiv
2. Densitatea populației României pe județe, la un moment dat, reprezintă o
 - a) distribuție de timp de momente
 - b) distribuție de timp de intervale
 - c) distribuție de spațiu
 - d) distribuție calitativă
3. Populația României înregistrată la 7 ianuarie 1992 și 18 martie 2002, reprezintă o
 - a) distribuție de timp de momente
 - b) distribuție de timp de intervale
 - c) distribuție de spațiu
 - d) distribuție calitativă
4. Puterea instalată a grupurilor electrogene din România, în anii 1990-2004, reprezintă o
 - a) distribuție de timp de momente
 - b) distribuție de timp de intervale
 - c) distribuție de spațiu
 - d) distribuție calitativă



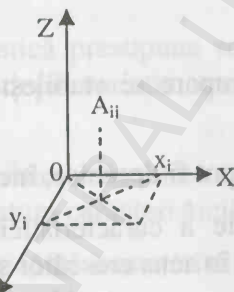
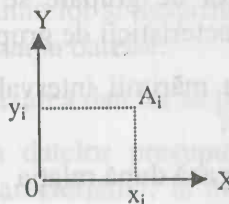
3.8. Răspunsuri corecte

3.6. Alegeți răspunsul corect

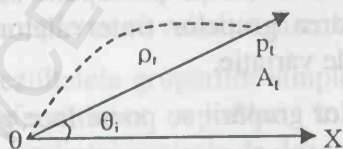
1. c) Sistematizarea datelor înregistrate într-un proces de observare statistică reprezintă procesul de obținere a distribuției statistice.
2. a) și c) Sistematizarea datelor presupune un set de operații, și anume: strângerea formularelor și desprinderea din formulare a datelor individuale; sortarea și totalizarea datelor.
3. b) și c) Sistematizarea datelor se poate realiza prin centralizare și grupare.
4. a) Centralizarea datelor presupune totalizarea unităților statistice sau a valorilor unei caracteristici, la nivelul grupelor tipice sau al colectivității observate.
5. b) Gruparea statistică presupune separarea unităților unei colectivități pe grupe omogene, după variația uneia sau a mai multor caracteristici de grupare.
6. a) și c) Gruparea datelor se poate efectua după numărul caracteristicilor de grupare și după natura caracteristicilor de grupare.
7. b) și c) După numărul caracteristicilor de grupare se pot obține grupări simple și grupări combinate.
8. a) Grupările simple au ca rezultat distribuții unidimensionale.
9. a), b) și c) Rezultatele grupărilor simple sunt folosite pentru caracterizarea structurii, diferențierea tipurilor în cadrul colectivității și pentru determinarea tendinței centrale de dezvoltare a fenomenelor.
10. b) și c) Grupările combinate au ca rezultat distribuții bidimensionale și distribuții multidimensionale.
11. a) și b) Grupările combinate sunt folosite pentru studierea legăturilor statistice și studierea influenței factorilor asupra variației.
12. a), b) și d) După natura caracteristicii de grupare, se pot distinge grupări după o variabilă de timp, grupări după o variabilă de spațiu și grupări după o variabilă de calitate.

13. a) și b) Grupările după o caracteristică calitativă atributivă au ca rezultat nomenclatoare statistice și clasificări statistice.
14. a) și c) Grupările după o caracteristică calitativă numerică au ca rezultat distribuții după o variabilă discretă și distribuții după o variabilă continuă.
15. a) Rezolvarea problemelor privind precizarea scopului grupării, alegerea variabilelor de grupare, stabilirea numărului de grupe, stabilirea mărimii intervalelor de grupare, delimitarea grupelor de variație și separarea unităților pe intervale de grupare reprezintă probleme ale grupării statistice.
16. a) și b) Scopul grupării constă în analiza directă, în cazul grupărilor tipologice, și sistematizarea materialului factual, în vederea prelucrării.
17. b) și c) Alegerea variabilei/variabilelor de grupare se face în funcție de scopul grupării și de esențialitate a caracteristicii de grupare.
18. a) Stabilirea numărului de grupe și a mărimii intervalelor de grupare se realizează în funcție de scopul grupării.
19. b) și c) Numărul de grupe (k) se calculează după relația
- $$k = A/l \quad \text{și}$$
- $$k = 1 + 3.322 \log n$$
20. a) și c) Mărimea intervalelor de grupare se stabilește conform relațiilor:
- $$l = A/k \quad \text{și} \quad l = A/(1 + 3.322 \log n).$$
21. a), b), c) și d) Intervalele de grupare pot fi deschise, închise, egale, inegale.
22. c) Gruparea pe variante de variație a caracteristicii/caracteristicilor de grupare presupune ordonarea datelor în sens crescător sau descrescător.
23. a), b) și c) Gruparea pe intervale de variație presupune stabilirea mărimii intervalului de variație, delimitarea grupelor (intervalelor) de variație și separarea unităților pe intervale de variație.
24. a), b) și c) Prezentarea rezultatelor grupării se poate face prin serii, tabele, grafice.
25. a), b), c) și d) Elementele unui tabel statistic sunt titlul general și titlurile interioare, rubricile, legenda, notele, sursa.
26. b) și c) Tabelele statistice sunt, de regulă, tabele simple și tabele cu dublă intrare.
27. a) și d) Tabelele cu dublă intrare sunt tabele de corelație, în cazul variabilelor numerice și tabele de asociere, în cazul variabilelor nominale.

28. c) Reprezentarea grafică este o metodă de prezentare vizuală a datelor, într-un sistem de coordonate dat.
29. a), b), c), d) și e) Elementele unui grafic sunt titlul, axele de coordonate, scara, rețeaua, diagrama și legenda.
30. a) și c) Axele de coordonate pentru construirea unui grafic statistic pot fi axe în sistem de coordonate rectangulare sau axe în sistem de coordonate polare.
31. a) și b) Reprezentarea unui punct într-un sistem de coordonate rectangulare este dată în figurile:



32. c) Reprezentarea unui punct într-un sistem de coordonate polare este dată în figura:



θ_i : unghi polar (în grade sau radiani)
 $(OA_i = \rho_i)$: raza vectoare

33. a) și c) Scara de reprezentare grafică poate fi uniformă (aritmetică) sau neuniformă (gaussiană).

34. a) Trecerea de la o scară la alta presupune realizarea unei corespondențe dintre o progresie geometrică de rație 10 și o progresie aritmetică de rație 1 (Bazele 1 și 0 sunt corespondente).

Numere	0,001	0,01	0,1	1	10	100	1000
Logaritmi zecimali	-3	-2	-1	0	1	2	3

*Date convenționale

35. b) Corespondența dintre orice număr 10^n (n oarecare) și logaritmul său este dată de relația $\log(10^n) = n$.
36. c) Prin logaritmare, modelele de trend de tip putere se liniarizează.
37. a) Tipul de reprezentare grafică se alege în funcție de natura seriilor (distribuțiilor) statistice.
38. a) Pentru reprezentarea grafică a unei serii calitative unidimensionale cu variabila exprimată numeric sunt folosite histograma, poligonul frecvențelor, curba frecvențelor, curba frecvențelor cumulate.
39. b) Pentru reprezentarea grafică a unei serii calitative unidimensionale cu variabila exprimată atributiv sunt folosite diagramele de structură.
40. c) Pentru reprezentarea grafică a unei serii bidimensionale cu ambele variabile exprimate cifric sunt folosite norul de puncte, diagramele prin paralelipipede, suprafața poliedrală.
41. c) Pentru reprezentarea grafică a unei serii bidimensionale cu ambele variabile nominale exprimate atributiv sunt folosite diagramele de structură construite în același plan.
42. a) Pentru reprezentarea grafică a unei serii bidimensionale cu o variabilă calitativă exprimată atributiv și o variabilă calitativă exprimată cifric sunt folosite diagrame tip „piramida vârstelor“.
43. b) și c) O distribuție unidimensională se prezintă sub forma

$$X : \begin{Bmatrix} x_1 & x_2 & \dots & x_i & \dots & x_m \\ n_1 & n_2 & \dots & n_i & \dots & n_m \end{Bmatrix}, i = \overline{1, m}$$

sau

$$X : (x_i), i = \overline{1, n}$$

44. a) O distribuție bidimensională se prezintă sub forma

$$X, Y : \{x_i, y_j, n_{ij}\}, i = \overline{1, m}, j = \overline{1, p}$$

45. a) și c) O distribuție marginală se prezintă sub forma

$$\{x_i, n_{i\cdot}\}, i = \overline{1, m} \quad \text{sau} \quad \{y_j, n_{\cdot j}\}, j = \overline{1, p}$$

46. b) și c) Distribuțiile condiționate se prezintă sub forma

$$(X|Y = y_j) : \{x_i, n_{ij}\}, i = \overline{1, m}, j - \text{valoare fixă}$$

$$(Y|X = x_i) : \{y_j, n_{ij}\}, j = \overline{1, p}, i - \text{valoare fixă}$$

47. a) și c) Frecvențele relative marginale sunt definite de relațiile

$$f_{i\cdot} = \frac{n_{i\cdot}}{n_{\cdot\cdot}} \quad f_{\cdot j} = \frac{n_{\cdot j}}{n_{\cdot\cdot}}$$

48. a) Frecvențele relative parțiale sunt definite de relația

$$f_{ij} = \frac{n_{ij}}{n_{\cdot\cdot}}$$

49. a) și b) Frecvențele relative condiționate sunt definite de relațiile

$$f_{ij|j} = \frac{n_{ij}}{n_{\cdot j}}, j - \text{valoare fixă}, i = \overline{1, m}$$

$$f_{j|i} = \frac{n_{ij}}{n_{i\cdot}}, i - \text{valoare fixă}, j = \overline{1, p}$$

3.7. Recunoașteți exemplul corect

1. d) Distribuția populației ocupate a unei țări pe sectoare de activitate, la un moment dat, este un exemplu de distribuție calitativă cu variabila exprimată atributiv.
2. c) Densitatea populației României pe județe, la un moment dat, este un exemplu de distribuție de spațiu.
3. a) Populația României înregistrată la 7 ianuarie 1992 și la 18 martie 2002 este un exemplu de distribuție de timp de momente.
4. b) Puterea instalată a grupurilor electrogene din România, în anii 1990-2004, este un exemplu de distribuție de timp de intervale.



3.9. Probleme

1. Distribuția a n familii dintr-un bloc, după numărul de copii (X), la un moment dat, se prezintă astfel:

x_i	0	1	2	3	4	5	6	7
n_i	6	18	23	20	14	6	2	1

*Date convenționale

Să se calculeze numărul de familii care au cel mult 4 copii.

Aflarea numărului de familii care au cel mult 4 copii presupune calculul frecvenței absolute cumulate.

Frecvența absolută cumulată crescător se calculează după relația:

$$N_i = n_1 + n_2 + \dots + n_i = \sum_{h=1}^i n_h$$

$$N_1 = 6$$

$$N_2 = 6 + 18 = 24$$

$\vdots$

$$N_5 = 6 + 18 + 23 + 20 + 14 = 81$$

Interpretare: Un număr de 81 familii din colectivitatea considerată au cel mult 4 copii.

2. Distribuția familiilor dintr-un bloc, după numărul de copii, la un moment dat, se prezintă astfel:

x_j	0	1	2	3	4	5	6	7
n_j	6	18	23	20	14	6	2	1

*Date convenționale

Să se calculeze ponderea familiilor care au cel mult 4 copii.

Aflarea ponderii familiilor care au cel mult 4 copii presupune calculul frecvenței relative cumulate.

Frecvența relativă cumulată crescător se calculează după relația:

$$F_i = F_{i-1} + f_i$$

$$F_i = \sum_{h=1}^i f_h$$

$$f_i = \frac{n_i}{\sum n_i} \cdot 100$$

$$f_1 = \frac{6}{90} \cdot 100 = 6,66\%$$

:

$$F_5 = 6,66\% + 20\% + 25,55\% + 22,22\% + 15,56\% = 90\%$$

Interpretare: Ponderea familiilor din colectivitatea considerată care au cel mult 4 copii este de 90%.

3. Distribuția pe grupe de vârstă a unui eșantion de 100 persoane dintr-un cartier de locuințe, la un moment dat, este

Grupa de vârstă (ani)	-20	20-30	30-60	60-80	80-100
Efectivul	17	61	22	15	18

*Date convenționale

Să se calculeze limita inferioară a primului interval de variație.

Exemplul dat prezintă o distribuție pe intervale inegale. În acest caz, limita inferioară a primului interval de variație se închide, luând în considerare mărimea intervalului următor,

$$d_i = x_i - x_{i-1}$$

Mărimea intervalului este :

$d_i = 30 - 20 = 10$. Prin urmare, primul interval are ca limită inferioară 10, adică (20-10).

4. Distribuția pe grupe de vârstă a unui eșantion de 100 persoane dintr-un cartier de locuințe este prezentată astfel:

Grupa de vârstă (ani)	10-15	15-60	60 și peste
Efectivul	17	61	22

*Date convenționale

Să se calculeze limita superioară a ultimului interval de variație.

Exemplul dat prezintă o distribuție pe intervale inegale. În acest caz, limita superioară a ultimului interval de variație se închide, luând în considerare mărimea intervalului anterior:

$$d_i = x_i - x_{i-1}$$

Mărimea intervalului anterior intervalului căutat este: $d_i = 60 - 15 = 45$.

Ca urmare, limita superioară a ultimului interval este 105, adică $(60 + 45)$.

5. În urma înregistrării unui eșantion de 200 de persoane privind timpul de deplasare zilnică, de la domiciliu la locul de muncă și invers, și a sistematizării datelor s-a obținut distribuția :

Timpul de deplasare (minute)	-30	30-60	60-90	90-120	120 -
Numărul persoanelor	25	50	60	45	15

**Date convenționale*

Să se calculeze limita inferioară a primului interval de variație.

Exemplul dat prezintă o distribuție pe intervale egale. În acest caz, limita inferioară a primului interval de variație se închide, luând în considerare aceeași mărime a intervalului:

$$d_i = x_i - x_{i-1}$$

Mărimea intervalelor de variație este $d_i = 60 - 30 = 30$, $d_i = 90 - 60 = 30$...

Prin urmare, limita inferioară a primului interval este 0, adică $(30 - 30)$.

6. În urma înregistrării unui eșantion de 200 de persoane privind timpul de deplasare zilnică, de la domiciliu la locul de muncă și invers, și a sistematizării datelor s-a obținut distribuția prezentată, astfel:

Timpul de deplasare (minute)	-30	30-60	60-90	90-120	120 -
Numărul persoanelor	25	50	60	45	15

**Date convenționale*

Să se calculeze limita superioară a ultimului interval de variație.

Exemplul dat prezintă o distribuție pe intervale egale. În acest caz, limita superioară a ultimului interval de variație se închide luând

Mărimea intervalelor de variație este $d_i = 60 - 30 = 30$, $d_i = 90 - 60 = 30$...

Prin urmare, limita superioară a ultimului interval este 150, adică

în considerație aceeași mărime a | $(120+30)$.
intervalului:

$$d_i = x_i - x_{i-1}$$

7. Distribuția unui eșantion de 200 de persoane după timpul de deplasare zilnică de la domiciliu la locul de muncă și invers este

Timpul de deplasare (minute)	- 30	30-60	60-90	90-120	120 și peste
Numărul persoanelor	25	50	60	45	15

*Date convenționale

Să se reprezinte grafic distribuția după timpul de deplasare pentru eșantionul observat, folosind histograma și curba frecvențelor.

Reprezentarea grafică a distribuției după timpul de deplasare pentru eșantionul observat, folosind histograma și curba frecvențelor, este realizată în figura 3.5.

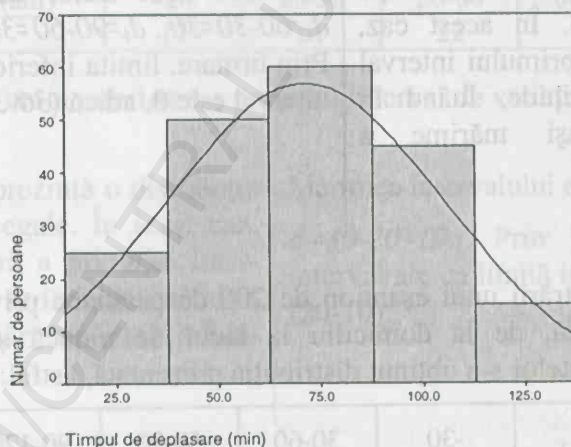


Fig. 3.5. Distribuția după timpul de deplasare pentru eșantionul observat

8. Distribuția pe grupe de vârstă a unui eșantion de indivizi dintr-o localitate, observați la un moment dat, se prezintă astfel :

Grupa de vârstă	Tânăr	Adult	În vârstă
Efectivul	17	61	22

*Date convenționale

Să se reprezinte grafic structura pe grupe de vârstă pentru eșantionul observat.

Reprezentarea grafică a structurii pe grupe de vârstă pentru eșantionul observat este realizată în figura 3.6.

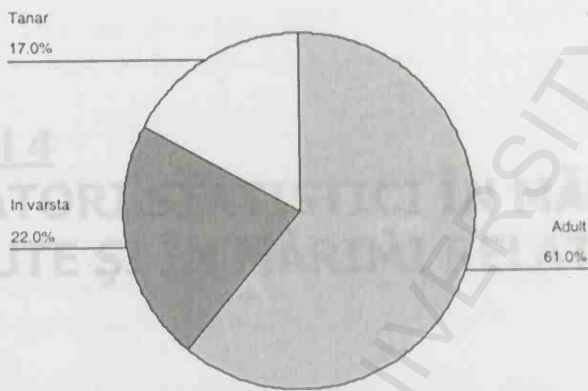


Fig. 3.6. Structura pe grupe de vârstă a eșantionului observat

Capitolul 4

INDICATORI STATISTICI ÎN MĂRIMI ABSOLUTE ȘI ÎN MĂRIMI RELATIVE

- ◆ Indicatori în mărimi absolute (indicatori de nivel, indicatori ai variației absolute)
- ◆ Indicatori în mărimi relative (mărimi relative de structură, mărimi relative de corespondență, mărimi relative de intensitate)
- ◆ Calculul mărimilor relative de structură în SPSS
- ◆ Alegeți răspunsul corect
- ◆ Recunoașteți exemplul corect
- ◆ Răspunsuri corecte
- ◆ Probleme

4.1. Indicatori în mărimi absolute

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 95-97)

- **Indicatori de nivel**
 - Indicatorii de nivel exprimă *aspectul cantitativ al fenomenelor, volumul unui ansamblu sau valoarea unei caracteristici*
 - *Indicatori individuali și indicatori sintetici*
- **Indicatori ai variației absolute**
 - Indicatorii variației absolute *se obțin prin compararea, sub formă de scădere, a două nivele ale aceluiași indicator*
 - *Indicatori individuali și indicatori sintetici*

4.2. Indicatori în mărimi relative

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 97-104)

- **Probleme ale folosirii mărimilor relative**
 - *Alegerea bazei de comparare*
 - *Asigurarea comparabilității datelor*
 - *Alegerea formei de exprimare a mărimilor relative*
- **Tipuri de mărimi relative**
 - *Mărimile relative de structură* exprimă raportul dintre parte și întreg și arată ponderea părții (grupe) față de totalul colectivității.
 - *Mărimile relative de intensitate* evidențiază gradul, intensitatea de răspândire a unui fenomen, în raport cu variabila la care se raportează.
 - *Mărimile relative ale dinamicii* exprimă variația relativă a unei variabile, X , între două momente/perioade considerate.

4.3. Calculul mărimilor relative de structură în SPSS

- Demersul în SPSS
 - Din meniul *Analyse*, se alege opțiunea *Frequencies*
- Rezultatele în SPSS se prezintă astfel:

SEXUL

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	masculin	18	45,0	45,0	45,0
	feminin	22	55,0	55,0	100,0
	Total	40	100,0	100,0	



4.4. Alegeți răspunsul corect

1. Indicatorii statistici primari se exprimă
 - a) în mărimi relative
 - b) în mărimi absolute
 - c) în mărimi medii
2. Mărimile absolute sunt
 - a) valori definite independent de orice sistem de referință
 - b) valori exprimate în unități naturale, natural-convenționale, unități valorice, unități de timp de muncă
 - c) valori exprimate în procente
3. Indicatorii statistici în mărimi absolute sunt
 - a) indicatori ai nivelului absolut
 - b) indicatori de structură
 - c) indicatori ai variației absolute
4. Indicatorii de nivel în mărimi absolute pot fi determinați ca
 - a) indicatori individuali
 - b) indicatori ai structurii
 - c) indicatori sintetici
5. Indicatorii individuali în mărimi absolute exprimă
 - a) rezultatul centralizării datelor pe grupe sau pe ansamblul colectivității ($\sum_i x_i$)
 - b) frecvența relativă corespunzătoare fiecărei unități statistice
 - c) valoarea „ x_i ” a caracteristicii X, observate la nivelul fiecărei unități statistice
6. Indicatorii sintetici în mărimi absolute reprezintă
 - a) rezultatul centralizării datelor pe grupe sau pe ansamblul colectivității ($\sum_i x_i$)
 - b) frecvența relativă corespunzătoare fiecărei unități statistice
 - c) valoarea „ x_i ” a caracteristicii X observate la nivelul fiecărei unități statistice

7. Însumarea directă este posibilă pentru
- a) elemente omogene
 - b) elemente eterogene
 - c) coeficienți de echivalență
8. Însumarea elementelor eterogene impune folosirea
- a) coeficienților de echivalență
 - b) coeficienților de variație
 - c) coeficienților de corelație
9. Indicatorii variației absolute se obțin prin
- a) compararea, sub formă de scădere, a două niveluri ale aceluiași indicator
 - b) $\Delta_{1/0}^x = x_{i1} - x_{i0}, i = \overline{1, n}$
 - c) $\Delta_{1/0}^x = \sum_i x_{i1} - \sum_i x_{i0}, i = \overline{1, n}$
 - d) compararea, sub formă de raport, a doi indicatori statistici
10. Indicatorii statistici în mărimi relative pot fi rezultați din
- a) compararea sub formă de raport a doi indicatori statistici
 - b) $f_i = \frac{n_i}{\sum_i n_i}, \sum_i f_i = 1, i = \overline{1, n}$
 - c) compararea sub formă de scădere a doi indicatori statistici
11. Indicatorii relativi arată
- a) cu cât s-a modificat un indicator în două momente T_0 și T_1 diferite
 - b) câte unități din indicatorul de la numărător revin la o unitate a indicatorului de la numitor
 - c) cu cât a crescut, în mărime absolută, nivelul unui indicator între două perioade comparate
12. Folosirea mărimilor relative presupune rezolvarea problemelor legate de
- a) alegerea bazei de comparare
 - b) asigurarea comparabilității datelor care formează raportul
 - c) alegerea formei de exprimare a mărimilor relative
13. Comparabilitatea datelor considerate în calculul mărimilor relative trebuie asigurată din punctul de vedere al
- a) metodologiei de culegere și prelucrare a datelor
 - b) formei de exprimare a indicatorilor
 - c) gradului de cuprindere a elementelor

14. Mărimile relative pot fi exprimate sub formă

- a) de %, ‰
- b) de coeficienți
- c) absolută

15. Mărimile statistice relative pot fi calculate ca

- a) mărimi de structură
- b) mărimi de corespondență (concordanță)
- c) mărimi de intensitate
- d) mărimi de variație

16. Mărimile relative de structură arată

- a) diferența dintre două niveluri ale aceluiași indicator
- b) suma dintre două niveluri ale aceluiași indicator
- c) raportul dintre parte și întreg

17. Mărimile relative de structură se calculează după relația

$$a) f_i = \frac{n_i}{\sum_i n_i}, i = \overline{1, m}$$

$$b) f_i = \frac{\sum_i n_i}{n_i}, i = \overline{1, m}$$

$$c) g_i = \frac{\sum_j x_{ij}}{\sum_i \sum_j x_{ij}}, i = \overline{1, m}, j = \overline{1, p}$$

$$d) \Delta_{1/0}^x = x_{i1} - x_{i0}, i = \overline{1, n}$$

18. Mărimile relative de corespondență măsoară

- a) câte unități dintr-o grupă revin la 100, respectiv 1000 de unități din altă grupă a unei colectivități
- b) cu cât s-a modificat nivelul unui indicator în două momente, T_0 și T_1 , diferite
- c) câte unități din indicatorul de la numărător revin la o unitate a indicatorului de la numitor

19. Mărimile relative de corespondență se calculează după relația

a) $\Delta_{i/0}^x = x_{i1} - x_{i0}, i = \overline{1, m}$

b) $K_{A/B} = \frac{X_A}{X_B}, K_{B/A} = \frac{X_B}{X_A}$

c) $f_i = \frac{n_i}{\sum_i n_i}, i = \overline{1, m}$

20. Mărimile relative de intensitate măsoară

- a) raportul dintre parte și întreg
- b) gradul, intensitatea de răspândire a unui fenomen, în raport cu variabila la care se raportează
- c) câte unități dintr-o grupă revin la 100, respectiv 1000 de unități din altă grupă a unei colectivități

21. Mărimile relative ale dinamicii măsoară

- a) variația relativă a unei variabile X între două momente/perioade considerate
- b) câte unități dintr-o grupă revin la 100, respectiv 1000 de unități din altă grupă a unei colectivități
- c) gradul, intensitatea de răspândire a unui fenomen, în raport cu variabila la care se raportează



4.5. Recunoașteți exemplul corect

1. O firmă are un efectiv de 120 angajați. Dintre aceștia 36 sunt ingineri. Ponderea inginerilor în totalul angajaților este de
 - a) 30%
 - b) 70%
 - c) 36%
2. Densitatea populației unei țări reprezintă o mărime relativă de
 - a) structură
 - b) intensitate
 - c) corespondență
3. Rata mortalității infantile reprezintă o mărime relativă de
 - a) structură
 - b) intensitate
 - c) corespondență



4.6. Răspunsuri corecte

4.4. Alegeți răspunsul corect

1. b) Indicatorii statistici primari se exprimă în mărimi absolute.
2. a) și b) Mărimile absolute sunt valori definite independent de orice sistem de referință, valori exprimate în unități naturale, natural-convenționale, unități valorice, unități de timp de muncă.
3. a) și c) Indicatorii statistici în mărimi absolute sunt indicatori ai nivelului absolut și indicatori ai variației absolute.
4. a) și c) Indicatorii de nivel în mărimi absolute pot fi determinați ca indicatori individuali și ca indicatori sintetici.
5. c) Indicatorii individuali în mărimi absolute exprimă valoarea „ x_i ” a caracteristicii X , observate la nivelul fiecărei unități statistice.
6. a) Indicatorii sintetici în mărimi absolute reprezintă rezultatul centralizării datelor pe grupe sau pe ansamblul colectivității ($\sum_i x_i$).
7. a) Însușirea directă este posibilă pentru elemente omogene.
8. a) Însușirea elementelor eterogene impune folosirea coeficienților de echivalență.
9. a), b) și c) Indicatorii variației absolute se obțin prin compararea, sub formă de scădere, a două niveluri ale aceluiași indicator, prin relația $\Delta_{1/0}^y = x_{i1} - x_{i0}$, $i = \overline{1, n}$, precum și prin relația $\Delta_{1/0}^x = \sum_i x_{i1} - \sum_i x_{i0}$, $i = \overline{1, n}$.
10. a) și b) Indicatorii statistici în mărimi relative pot fi rezultați din compararea sub formă de raport a doi indicatori statistici, precum și prin calculul frecvențelor relative, după relația: $f_i = \frac{n_i}{\sum_i n_i}$, $\sum_i f_i = 1$, $i = \overline{1, n}$.

11. b) Indicatorii relativi arată câte unități din indicatorul de la numărător revin la o unitate a indicatorului de la numitor.

12. a), b) și c) Folosirea mărimilor relative presupune rezolvarea problemelor legate de alegerea bazei de comparare, asigurarea comparabilității datelor care formează raportul și alegerea formei de exprimare a mărimilor relative.

13. a) și c) Comparabilitatea datelor considerate în calculul mărimilor relative trebuie asigurată din punctul de vedere al metodologiei de culegere și prelucrare a datelor și din punctul de vedere al gradului de cuprindere a elementelor.

14. a) și b) Mărimile relative pot fi exprimate sub formă de %, ‰, și de coeficienți.

15. a), b), c) și d) Mărimile statistice relative pot fi calculate ca mărimi de structură, mărimi de corespondență (concordanță), mărimi de intensitate și mărimi de variație.

16. c) Mărimile relative de structură arată raportul dintre parte și întreg.

17. a) și c) Mărimile relative de structură se calculează după relațiile

$$f_i = \frac{n_i}{\sum_i n_i}, i = \overline{1, m}; g_i = \frac{\sum_j x_{ij}}{\sum_i \sum_j x_{ij}}, i = \overline{1, m}, j = \overline{1, p}.$$

18. a) Mărimile relative de corespondență măsoară câte unități dintr-o grupă revin la 100, respectiv 1000 de unități din altă grupă a unei colectivități.

19. b) Mărimile relative de corespondență se calculează după relația

$$K_{A/B} = \frac{X_A}{X_B}, K_{B/A} = \frac{X_B}{X_A}.$$

20. b) Mărimile relative de intensitate măsoară gradul, intensitatea de răspândire a unui fenomen, în raport cu variabila la care se raportează.

21. a) Mărimile relative ale dinamicii măsoară variația relativă a unei variabile X între două momente/perioade considerate.

Relație de calcul:

4.5. Recunoașteți exemplul corect

- a) O firmă are un efectiv de 120 angajați. Dintre aceștia 36 sunt ingineri. Ponderea inginerilor în totalul angajaților este: $f_i = \frac{36}{120} \cdot 100 = 30\%$.
- b) Densitatea populației unei țări reprezintă o mărime relativă de intensitate.
- b) Rata mortalității infantile reprezintă o mărime relativă de intensitate.

Sectorul de activitate		1993	2001
Industria		3080	2717
Construcții		374	370
Agricultură și silvicultură		3674	2492
Alte ramuri		2844	2700
Total		10000	8553



4.7. Probleme

1. Populația ocupată (mii persoane) pe sectoare de activitate în România, în anii 1993 și 2001, se prezintă astfel:

Sectorul de activitate	1993	2001
Industrie	3030	2017
Construcții	574	340
Agricultură și silvicultură	3614	3498
Alte ramuri	2844	2708
Total	10062	8563

*Sursa: Anuarul statistic al României, 1994, C.N.S., p. 158, 2002, p. 94

Să se calculeze ponderea populației ocupate în sectorul de construcții, în anul 1993.

Relație de calcul:

$$f_i = \frac{n_i}{\sum n_i} \cdot 100$$

Ponderea populației ocupate în sectorul de construcții, în anul 1993, este:

$$f_2 = \frac{574}{10062} \cdot 100 = 5,7\%$$

Interpretare: Ponderea populației ocupate în sectorul de construcții, în anul 1993, a fost de 5,7%.

2. Populația ocupată (mii persoane) pe sectoare de activitate în România, în anii 1993 și 2001, se prezintă astfel:

Sectorul de activitate	1993	2001
Industrie	3030	2017
Construcții	574	340
Agricultură și silvicultură	3614	3498
Alte ramuri	2844	2708
Total	10062	8563

*Sursa: Anuarul statistic al României, 1994, C.N.S., p. 158, 2002, p. 94

Să se calculeze ponderea populației ocupate în sectorul de construcții, în anul 2001.

Relație de calcul:

$$f_i = \frac{n_i}{\sum_i n_i} \cdot 100$$

$$f_2 = \frac{340}{8563} \cdot 100 = 3,97\%$$

Interpretare: Ponderea populației ocupate în sectorul de construcții, în anul 2001, este de 3,97%.

3. Populația ocupată (mii persoane) pe sectoare de activitate în România, în anii 1993 și 2001, se prezintă astfel:

Sectorul de activitate	1993	2001
Industrie	3030	2017
Construcții	574	340
Agricultură și silvicultură	3614	3498
Alte ramuri	2844	2708
Total	10062	8563

*Sursa: Anuarul statistic al României, 1994, C.N.S., p. 158, 2002, p. 94

Să se afle cu cât s-a modificat structura populației ocupate în industrie, în anul 2001 față de 1993, în mărime absolută.

Relații de calcul:

$$f_i = \frac{n_i}{\sum_i n_i} \cdot 100$$

$$\Delta(\pm) = f_{i1} - f_{i0}$$

$$f_{i0} = \frac{3030}{10062} \cdot 100 = 30,11\%$$

$$f_{i1} = \frac{2017}{8563} \cdot 100 = 23,55\%$$

$$\Delta(\pm) = f_{i1} - f_{i0} = 23,55 - 30,11 = -6,56\%$$

Interpretare: Structura populației ocupate în industrie a scăzut, în anul 2001 față de 1993, cu 6,56%.

Capitolul 5

INDICATORI AI TENDINȚEI CENTRALE

- ◆ Probleme generale ale mărimilor medii
- ◆ Media aritmetică
- ◆ Medii cu aplicații speciale (media geometrică, media armonică, media pătratică)
- ◆ Modul sau dominanta
- ◆ Mediana
- ◆ Mediala
- ◆ Relații între valorile tendinței centrale
- ◆ Alegerea celei mai semnificative valori centrale
- ◆ Calculul indicatorilor tendinței centrale în SPSS
- ◆ Alegeți răspunsul corect
- ◆ Recunoașteți exemplul corect
- ◆ Răspunsuri corecte
- ◆ Probleme

5.1. Probleme generale ale mărimilor medii

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 105-107)

Mediile sunt mărimi statistice care exprimă, în mod sintetic și generalizant, ceea ce este normal, esențial, tipic pentru unitățile unei colectivități distribuite după o anumită caracteristică.

5.2. Media aritmetică

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 107-116)

- **Definiție**

Media aritmetică, categorie metodologică fundamentală a statisticii, reprezintă valoarea pe care ar purta-o fiecare unitate statistică dacă distribuția ar fi omogenă.

- **Mod de calcul**

Media aritmetică se calculează ca sumă a valorilor individuale, x_i , $i = \overline{1, n}$, împărțită la numărul de observări.

Media simplă:
$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i,$$

unde: $\sum_{i=1}^n x_i = x_1 + x_2 + \dots + x_n$, iar n este volumul colectivității.

Media ponderată:
$$\bar{x} = \frac{1}{\sum_{i=1}^m n_i} \cdot \sum_{i=1}^m x_i n_i, \text{ când } n_1 \neq n_2 \neq \dots \neq n_i;$$

$$\bar{x} = \sum_{i=1}^m x_i f_i, \quad \sum_{i=1}^m f_i = 1$$

Pentru serii de intervale $X : (J_i, n_i)$, cu $J_i = (x_{i-1}, x_i)$, $i = \overline{1, k}$

$$\bar{x} = \frac{1}{\sum_{i=1}^k n_i} \cdot \sum_{i=1}^k x'_i n_i = \sum_{i=1}^k x'_i f_i, \text{ unde } x'_i = \frac{x_{i-1} + x_i}{2}.$$

- **Clasificarea mărimilor medii**

După rolul în analiza statistică, distingem:

- *mărimi medii fundamentale* (media aritmetică, modul, mediana)
- *mărimi medii cu aplicații speciale* (media geometrică, media armonică, media pătratică, media progresivă, mediana, media cronologică, medii mobile etc.).

După modul de obținere, mărimile medii pot fi:

- *mărimi medii de calcul* (media aritmetică, media geometrică, media armonică etc.)
- *mărimi medii de poziție* (modul, mediana, mediana).

- **Proprietăți ale mediei aritmetice**

- **Calculul simplificat al mediei**

- **Media unei caracteristici alternative**

Tabelul 5.1. Distribuția de frecvență pentru o caracteristică alternativă

Valori ale caracteristicii (x_i)	Frecvența de apariție	
	Efectiv (n_i)	Pondere (f_i)
Da (1)	Unități care posedă caracteristica n_1	$p = \frac{n_1}{n}$
Nu (0)	Unități care nu posedă caracteristica ($n - n_1$)	$q = \frac{n - n_1}{n}$
Total	n	$p + q = 1$

Media pentru o caracteristică alternativă se află plecând de la media aritmetică ponderată. Astfel, folosind notațiile din tabelul 5.1, se poate scrie:

$$\bar{x} = \frac{\sum x_i n_i}{\sum n_i} = \frac{1 \cdot n_1 + 0 \cdot (n - n_1)}{n} = \frac{n_1}{n} = p.$$

5.3. Medii cu aplicație specială

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 116-121)

- Media geometrică ($\bar{x}_g$)
 - Media geometrică simplă
 - Media geometrică ponderată
- Media armonică ($\bar{x}_h$)
- Media pătratică ($\bar{x}_p$)
 - Media pătratică simplă
 - Media pătratică ponderată
- Media de ordin „r“
- Relația între mediile calculate pentru aceeași serie

$$\bar{x}_h < \bar{x}_g < \bar{x} < \bar{x}_p$$

5.4. Modul (M_0) sau Dominanta (D_0)

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 121-129)

- Definiție
 - Modul este valoarea caracteristicii care este cel mai frecvent observată într-o distribuție, adică valoarea ce corespunde frecvenței dominante. Este o mărime medie fundamentală, de poziție.
 - Se determină numai în cazul seriilor cu frecvențe diferite ($n_1 \neq n_2 \neq \dots \neq n_i$).

• **Determinarea grafică a modului. Cazul unei variabile discrete**

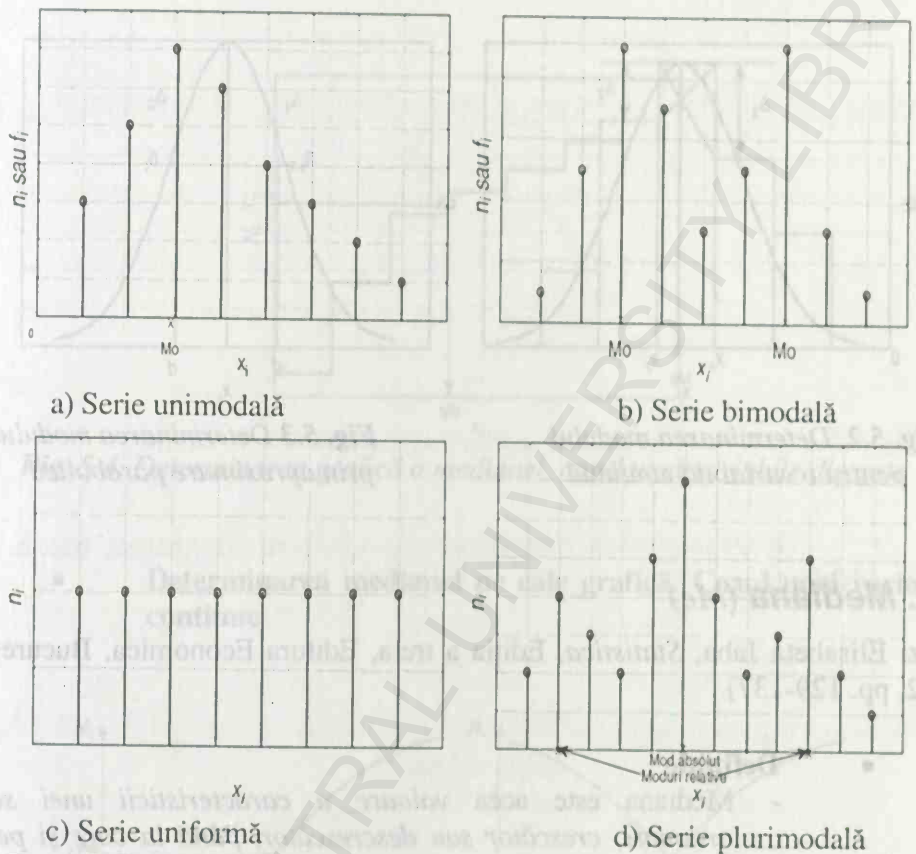


Fig. 5.1. Determinarea grafică a modului, în cazul unei variabile discrete

- **Determinarea grafică a modului. Cazul unei variabile continue**

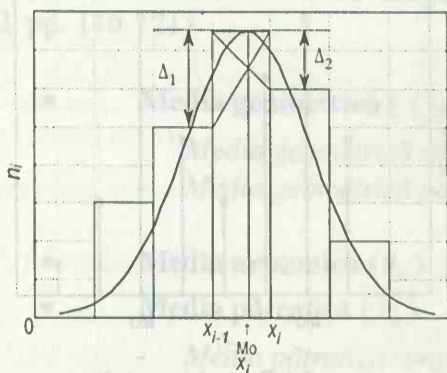


Fig. 5.2. Determinarea modului pentru o variabilă continuă

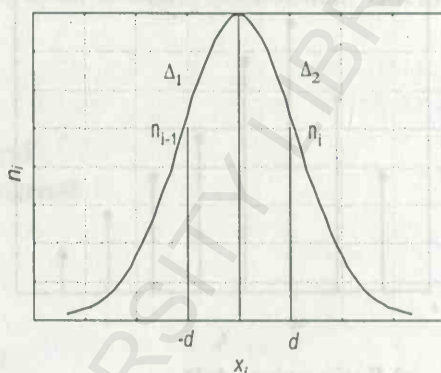


Fig. 5.3 Determinarea modului prin aproximare parabolică

5.5. Mediana (Me)

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 129-137)

Definiție

- Mediana este acea valoare a caracteristicii unei serii ordonate, crescător sau descrescător, până la care și peste care sunt distribuite în număr egal unitățile colectivității observate: jumătate din unități au valori mai mari decât mediana și jumătate au valori mai mici. Este o mărime medie fundamentală, de poziție.

Mod de calcul

- Pentru o serie simplă cu număr impar de termeni
- Pentru o serie simplă cu număr par de termeni
- Pentru o serie cu frecvență

- **Determinarea medianeii pe cale grafică. Cazul unei variabile discrete**

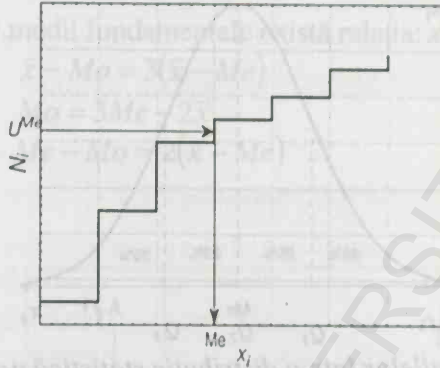


Fig. 5.4. Determinarea medianeii pe cale grafică, cazul unei variabile discrete

- **Determinarea medianeii pe cale grafică. Cazul unei variabile continue**

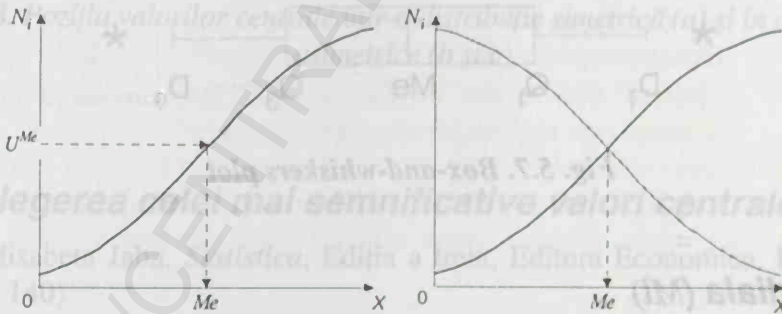


Fig. 5.5. Determinarea medianeii pe cale grafică, cazul unei variabile continue

- **Generalizarea medianei**
 - *Quantile*: quartile, decile, centile

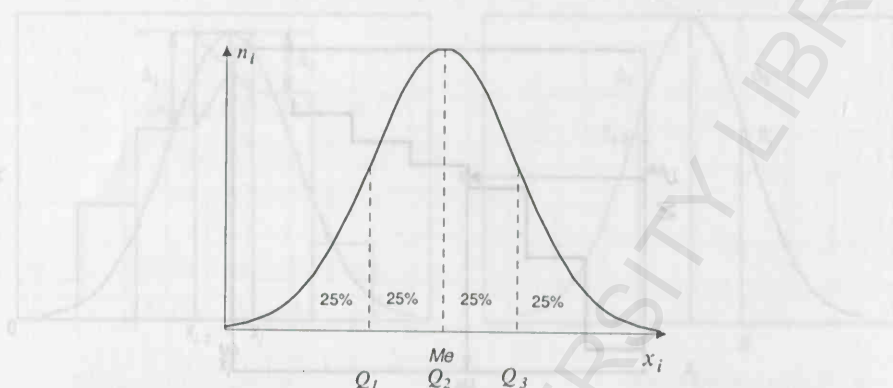


Fig. 5.6. Locul quartilelor într-o distribuție statistică normală

- **Reprezentarea quartilelor cu ajutorul diagramei box-and-whiskers**



Fig. 5.7. Box-and-whiskers plot

5.6. Mediala (Ml)

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 137-139)

- **Definiție**
 - *Mediala este un indicator de poziție egal cu acel nivel al caracteristicii (x_i) care împarte suma termenilor seriei ($\sum x_i n_i$) în două părți egale.*

- **Mod de calcul**

5.7. Relații între valorile tendinței centrale

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 139-140)

Între mărimile medii fundamentale există relația: $\bar{x} = Mo = Me$, respectiv:

$$\bar{x} - Mo = 3(\bar{x} - Me)$$

$$Mo = 3Me - 2\bar{x}$$

$$Me - Mo = 2(\bar{x} - Me)$$

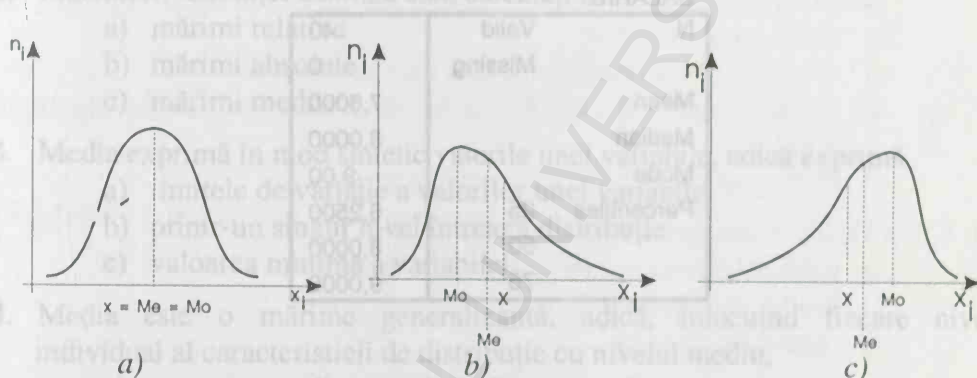


Fig. 5.8. Poziția valorilor centrale într-o distribuție simetrică (a) și în distribuții asimetrice (b și c)

5.8. Alegerea celei mai semnificative valori centrale

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, p. 140)

5.9. Calculul indicatorilor tendinței centrale în SPSS

• Demersul în SPSS

- Meniul *Analyse*, opțiunea *Descriptives Statistics*, comanda *Frequencies*

- În fereastra *Frequencies*, se alege *Statistics: Percentile Values* și *Central Tendency* (Mean, Median, Mode)

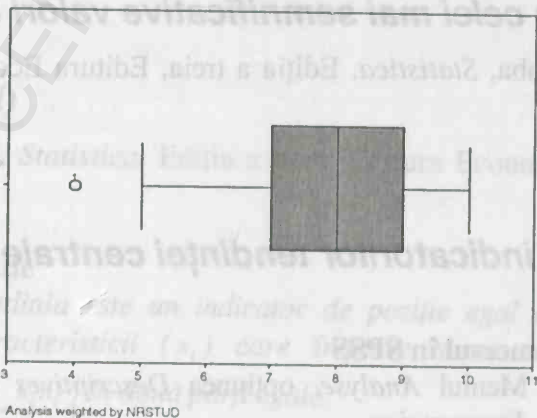
Rezultatele în SPSS

- Rezultatele în SPSS prezintă valorile indicatorilor tendinței centrale: media (Mean), mediana (Median), modul (Mode), quartilele unu, doi și trei (Percentiles).

Statistics

SALARIU		
N	Valid	40
	Missing	0
Mean		7,6000
Median		8,0000
Mode		9,00
Percentiles	25	6,2500
	50	8,0000
	75	9,0000

- Diagrama *box-and-whiskers* este reprezentată astfel:
Construirea diagramei *box-and-whiskers* presupune selectarea opțiunii *Charts* din meniul *Graphs*. Reprezentarea grafică a acesteia este realizată în figura de mai jos:





5.10. Alegeți răspunsul corect

1. Rolul indicatorilor tendinței centrale ai unei distribuții statistice este de a
 - a) sintetiza datele
 - b) calcula mărimile relative
 - c) aprecia gradul de dispersie al colectivității
2. Indicatorii tendinței centrale sunt calculați în
 - a) mărimi relative
 - b) mărimi absolute
 - c) mărimi medii
3. Media exprimă în mod sintetic valorile unei variabile, adică exprimă
 - a) limitele de variație a valorilor unei variabile
 - b) printr-un singur nivel întreaga distribuție
 - c) valoarea minimă a variabilei
4. Media este o mărime generalizantă, adică, înlocuind fiecare nivel individual al caracteristicii de distribuție cu nivelul mediu,
 - a) suma termenilor seriei rămâne aceeași
 - b) suma termenilor seriei se modifică
 - c) suma termenilor seriei este nulă
5. Media este o mărime abstractă, deoarece
 - a) exprimă în mod sintetic valorile centrale ale caracteristicii
 - b) exprimă în mod sintetic toate unitățile statistice ale distribuției
 - c) exprimă în mod sintetic toate valorile individuale ale unei caracteristici
6. Mărimile medii exprimă
 - a) ceea ce este normal într-o distribuție
 - b) structura colectivității analizate
 - c) ceea ce este tipic, legic într-o distribuție
7. Nivelul mediu al unei distribuții este un nivel normal, adică
 - a) este nivelul purtat de toate unitățile colectivității
 - b) este nivelul purtat de majoritatea variabilelor statistice
 - c) este nivelul purtat de majoritatea unităților unei colectivități

8. Media exprimă ceea ce este legic, adică
- nivelul datorat acțiunii factorilor esențiali
 - nivelul datorat acțiunii factorilor întâmplători
 - abaterile în plus și în minus de la nivelul mediu, datorate acțiunii factorilor aleatori, se compensează
9. Condițiile de calitate ale unei mărimi medii au fost sintetizate de
- Pearson (1945)
 - Fischer (1975)
 - Yule (1945)
10. Conform condițiilor de calitate, media trebuie
- să fie definită obiectiv
 - să fie reprezentativă
 - să aibă o semnificație concretă, ușor de observat
 - să fie simplu de calculat și să se preteze la calcule algebrice ulterioare
 - să fie puțin sensibilă la fluctuațiile de selecție
11. Criteriile de clasificare a mărimilor medii sunt
- rolul mărimilor medii în analiza statistică
 - modul de exprimare
 - modul de obținere
12. După rolul pe care îl au în analiza statistică, mărimile medii sunt
- mărimi medii fundamentale ($\bar{x}$, Do, Me)
 - mărimi medii cu aplicație specială ($\bar{x}$, $\bar{x}_g$, $\bar{x}_h$, $\bar{x}_p$, mediana, media cronologică și mediile mobile)
 - mărimi medii de calcul ($\bar{x}$, $\bar{x}_g$, $\bar{x}_h$)
13. După modul de obținere, mărimile medii se clasifică în
- mărimi medii de calcul
 - mărimi medii fundamentale
 - mărimi medii de poziție
14. Mărimile medii de calcul se obțin prin efectuarea a două categorii de operații
- acumularea unităților statistice (fie prin însumare, fie prin produs)
 - acumularea termenilor seriei (fie prin însumare, fie prin produs)
 - revenirea la un nivel reprezentativ pentru toți termenii incluși în calcul

15. Mărimile medii de poziție se obțin

- a) prin depistarea termenului care ocupă poziția centrală într-o distribuție statistică
- b) prin acumularea termenilor seriei
- c) prin depistarea termenului care ocupă o poziție extremă într-o distribuție statistică

16. Mărimile medii de calcul pot fi calculate ca

- a) medii simple
- b) medii de poziție
- c) medii ponderate

17. Mediile simple se calculează pentru

- a) $X: \{x_i, n_i\}, n_1 = n_2 = \dots = n_i, i = \overline{1, n}$
- b) $X: \{x_i, n_i\}, n_1 \neq n_2 \neq \dots \neq n_i, i = \overline{1, m}$
- c) $X: \{x_i\}, i = \overline{1, n}$

18. Mediile ponderate se calculează pentru

- a) $X: \{x_i, n_i\}, n_1 = n_2 = \dots = n_i, i = \overline{1, n}$
- b) $X: \{x_i, n_i\}, n_1 \neq n_2 \neq \dots \neq n_i, i = \overline{1, m}$
- c) $X: \{x_i\}, i = \overline{1, n}$

19. Media aritmetică reprezintă

- a) valoarea pe care ar purta-o fiecare unitate statistică dacă distribuția ar fi eterogenă
- b) valoarea pe care ar purta-o fiecare unitate statistică dacă distribuția ar fi omogenă
- c) valoarea pe care ar purta-o fiecare variabilă statistică dacă distribuția ar fi omogenă

20. Media aritmetică simplă se calculează după relația

a)
$$\bar{x} = \frac{1}{\sum_i n_i} \sum_i x_i n_i$$

b)
$$\bar{x} = \frac{1}{\sum_i n_i} \sum_i x_i$$

c)
$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

21. Media aritmetică ponderată se calculează după relația

$$a) \bar{x} = \frac{1}{\sum_i n_i} \sum_i x_i n_i, i = \overline{1, m}$$

$$b) \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$c) \bar{x} = \sum_i x_i f_i, \sum_i f_i = 1, i = \overline{1, m}$$

22. În cazul seriilor de intervale de variație, media aritmetică se calculează după relația

$$a) \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$b) \bar{x} = \frac{1}{\sum_i n_i} \sum_i x'_i n_i, i = \overline{1, k}, x'_i = \frac{1}{2}(x_{i-1} + x_i)$$

$$c) \bar{x} = \sum_i x'_i f_i, \sum_i f_i = 1, i = \overline{1, m}, x'_i = \frac{1}{2}(x_{i-1} + x_i)$$

23. Media aritmetică este

- a) o mărime normală
- b) mărime internă
- c) mărime translativă

24. Relațiile $\sum_i (x_i - \bar{x}) = 0, \sum_i (x_i - \bar{x}) n_i = 0$ arată că media este

- a) o mărime normală
- b) mărime translativă
- c) mărime internă

25. Media sumei a două variabile aleatoare independente, X și Y, este egală cu:

- a) suma mediilor celor două variabile, adică $\overline{x + y} = \bar{x} + \bar{y}$
- b) produsul mediilor celor două variabile, adică $\overline{x + y} = \bar{x} \cdot \bar{y}$
- c) raportul mediilor celor două variabile, adică $\overline{x + y} = \frac{\bar{x}}{\bar{y}}$

26. Media produsului a două variabile aleatoare independente, X și Y, este egală cu:

- a) suma mediilor celor două variabile, adică $\overline{x \cdot y} = \bar{x} + \bar{y}$
- b) produsul mediilor celor două variabile, adică $\overline{x \cdot y} = \bar{x} \cdot \bar{y}$

c) raportul mediilor celor două variabile, adică $\overline{x \cdot y} = \frac{\bar{x}}{\bar{y}}$

27. Relațiile $\bar{x} = \bar{x}' + a$, $\bar{x} = \bar{x}'' \cdot k$, $\bar{x} = \frac{\sum_i x_i \frac{n_i}{c}}{\sum_i \frac{n_i}{c}}$, unde: $\bar{x}' = \frac{1}{n} \sum_i (x_i \pm a)n_i$, iar

$$\bar{x}'' = \frac{1}{n} \sum_i \frac{x_i}{k} n_i, \text{ arată că media este}$$

- a) o mărime normală
- b) mărime translativă
- c) mărime internă

28. Calculul simplificat al mediei se poate efectua după relația

$$\text{a) } \bar{x} = \frac{\sum_i \left(\frac{x_i - a}{k} \right)}{n} \cdot k + a, i = \overline{1, n}$$

$$\text{b) } \bar{x} = \frac{\sum_i \left(\frac{x_i - a}{k} \right)}{n} \cdot k - a, i = \overline{1, n}$$

$$\text{c) } \bar{x} = \frac{\sum_i \left(\frac{x_i - a}{k} \right) \cdot \frac{n_i}{c}}{\sum_i \frac{n_i}{c}} \cdot k + a, i = \overline{1, n}$$

29. Media unei caracteristici alternative se calculează după relația

$$\text{a) } \bar{x} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n n_i}$$

$$\text{b) } p = \frac{n_i}{n}$$

$$\text{c) } q = \frac{n - n_i}{n}$$

30. Media geometrică se calculează după relațiile

a) $\bar{x}_g = \sqrt[n]{\prod_{i=1}^n x_i}, n > 2$

b) $\bar{x}_g = \sqrt[n]{\prod_{i=1}^n x_i}$

c) $\bar{x}_g = \sqrt[n]{\prod_{i=1}^n x_i^{n_i}}$

31. Media geometrică a unei variabile Z, definită ca produs a două variabile, X și Y, este egală cu

a) raportul mediilor geometrice ale celor două variabile, adică $\bar{Z} = \bar{X}_g / \bar{Y}_g$

b) produsul mediilor geometrice ale celor două variabile, $\bar{Z}_g = \bar{X}_g \cdot \bar{Y}_g$

c) suma mediilor geometrice ale celor două variabile, $\bar{Z}_g = \bar{X}_g + \bar{Y}_g$

32. Media geometrică a unei variabile Z, definită ca raport al două variabile, X și Y, este egală cu

a) raportul mediilor geometrice ale celor două variabile, adică $\bar{Z} = \bar{X}_g / \bar{Y}_g$

b) produsul mediilor geometrice ale celor două variabile, $\bar{Z}_g = \bar{X}_g \cdot \bar{Y}_g$

c) suma mediilor geometrice ale celor două variabile, $\bar{Z}_g = \bar{X}_g + \bar{Y}_g$

33. Modul reprezintă

- a) valoarea caracteristicii purtată de majoritatea unităților populației
- b) valoarea caracteristicii cel mai frecvent observate într-o distribuție
- c) valoarea caracteristicii care împarte volumul colectivității în două părți egale

34. Mediana reprezintă

- a) valoarea caracteristicii purtată de majoritatea unităților populației
- b) valoarea caracteristicii cel mai frecvent observate într-o distribuție
- c) valoarea caracteristicii care împarte volumul colectivității în două părți egale

35. Unitatea mediană corespunde
- locului medianei în distribuție
 - acelei unități din populație care poartă nivelul mediu al caracteristicii
 - acelei unități din populație care poartă nivelul maxim al caracteristicii
36. Pentru o serie statistică dată, între mărimile medii de calcul există următoarele relații
- $\bar{x}_h < \bar{x}_g < \bar{x} < \bar{x}_p$
 - $\bar{x}_p > \bar{x} > \bar{x}_g > \bar{x}_h$
 - $\bar{x}_g > \bar{x}_p < \bar{x}_h > \bar{x}$
 - $\bar{x}_h > \bar{x}_p < \bar{x} < \bar{x}_g$
37. Mediana este o mărime medie
- de poziție
 - cu aplicație specială
 - fundamentală
 - de calcul
38. Modul se utilizează pentru caracterizarea
- tendenței centrale
 - gradului de asimetrie
 - gradului de boltire
 - gradului de concentrare
39. Quantilele sunt mărimi de poziție care împart colectivitatea în
- $n-1$ părți egale
 - 4 părți egale
 - 5 părți egale
 - n părți egale
40. Mediala este o mărime
- medie de poziție
 - care împarte colectivitatea în două părți egale
 - care împarte suma valorilor globale ale seriei în două părți egale
 - medie egală întotdeauna cu Me
41. Determinarea modului, în cazul seriilor de intervale inegale, presupune
- calculul frecvențelor reduse
 - reducerea lungimii intervalelor

c) calculul mijlocului intervalelor

42. Modul se află după relația

a) $Mo = x_{i-1} + d \cdot \frac{\Delta_1}{\Delta_1 + \Delta_2}$

b) $Mo = x_{i-1} - d \cdot \frac{\Delta_1}{\Delta_1 + \Delta_2}$

c) $Mo = x_{i-1} + d \cdot \frac{(n_i - n_{i-1})}{(n_i - n_{i-1}) + (n_i - n_{i+1})}$

43. Mediana se află după relația

a) $Me = x_{i-1} + d \cdot \frac{\sum n_i - N_{i-1}}{n_i}$

b) $Me = x_{i-1} + d \cdot \frac{U^{Me} - N_{i-1}}{n_i}$

c) $Me = x_{i-1} + d \cdot \frac{U^{Me} + N_{i-1}}{n_i}$



5.11. Recunoașteți exemplul corect

1. Considerăm rata inflației cu următoarele valori înregistrate trei ani consecutivi: 10% în primul an, 10% în al doilea an și 70% în al treilea an și o valoare de referință, la începutul perioadei studiate, $x_0 = 100 \text{ USD}$.

Rata medie a inflației $= \sqrt[3]{1,1 \cdot 1,1 \cdot 1,7} = 1,2718$ este

- a) o medie aritmetică
- b) o medie geometrică
- c) o medie armonică

2. Într-o întreprindere cu 100 de angajați, salariul angajaților urmează o lege normală, $X \sim N(3,4;4)$. Salariul mediu s-a calculat ca

- a) o medie aritmetică
- b) medie geometrică
- c) medie armonică

3. 25% din salariații unei întreprinderi au un salariu de cel mult 4 milioane lei. Această valoare reprezintă

- a) quartila unu
- b) mediana
- c) quartila trei

4. 75% din salariații unei întreprinderi au un salariu de cel mult 7 milioane lei. Această valoare reprezintă

- a) quartila unu
- b) mediana
- c) quartila trei

5. 50% din salariații unei întreprinderi au un salariu de cel mult 5 milioane lei. Această valoare reprezintă

- a) quartila unu
- b) mediana
- c) quartila trei

6. În distribuția de mai jos:

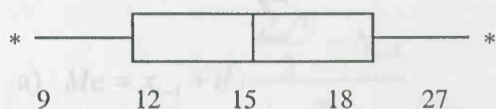
(x_{i-1}, x_i)	0-5	5-10	10-15	15-20
n_i	3	7	15	20

*Date convenționale

intervalul (15-20) reprezintă

- intervalul median
- intervalul modal
- intervalul dominant

7. Box-and-whiskers plot de mai jos



arată

- o valoare a quartilei unu egală cu 27
- o valoare a quartilei trei egală cu 18
- o valoare mediană egală cu 15



5.12. Răspunsuri corecte

5.10. Alegeți răspunsul corect

1. a) Rolul indicatorilor tendinței centrale ai unei distribuții statistice este de a sintetiza datele.
2. c) Indicatorii tendinței centrale sunt calculați în mărimi medii.
3. b) Media exprimă în mod sintetic valorile unei variabile, adică exprimă printr-un singur nivel întreaga distribuție.
4. a) Media este o mărime generalizantă, adică, înlocuind fiecare nivel individual al caracteristicii de distribuție cu nivelul mediu, suma termenilor seriei rămâne aceeași.
5. c) Media este o mărime abstractă, deoarece exprimă în mod sintetic toate valorile individuale ale unei distribuții.
6. a) și c) Mărimile medii exprimă ceea ce este normal într-o distribuție și ceea ce este tipic, legic într-o distribuție.
7. c) Nivelul mediu al unei distribuții este un nivel normal, adică este nivelul purtat de majoritatea unităților unei colectivități.
8. a) și c) Media exprimă ceea ce este legic, adică nivelul datorat acțiunii factorilor esențiali și abaterile în plus și în minus de la nivelul mediu, datorate acțiunii factorilor aleatori, se compensează.
9. c) Condițiile de calitate ale unei mărimi medii au fost sintetizate de Yule (1945).
10. a), b), c), d) și e) Conform condițiilor de calitate, media trebuie să fie definită obiectiv, să fie reprezentativă, să aibă o semnificație concretă, ușor de observat, să fie simplu de calculat și să se preteze la calcule algebrice ulterioare și să fie puțin sensibilă la fluctuațiile de selecție.
11. a) și c) Criteriile de clasificare a mărimilor medii sunt rolul mărimilor medii în analiza statistică și modul de obținere.

12. a) și b) După rolul pe care îl au în analiza statistică, mărimile medii sunt mărimi medii fundamentale ($\bar{x}$, Do, Me) și mărimi medii cu aplicație specială ($\bar{x}_g, \bar{x}_h, \bar{x}_p$, mediana, mediile cronologice și mediile mobile).
13. a) și c) După modul de obținere, mărimile medii se clasifică în mărimi medii de calcul și mărimi medii de poziție.
14. b) și c) Mărimile medii de calcul se obțin prin efectuarea a două categorii de operații: acumularea termenilor seriei (fie prin însumare, fie prin produs) și revenirea la un nivel reprezentativ pentru toți termenii incluși în calcul.
15. a) Mărimile medii de poziție se obțin prin depistarea termenului care ocupă poziția centrală într-o distribuție statistică.
16. a) și c) Mărimile medii de calcul pot fi calculate ca medii simple și medii ponderate.
17. a) și c) Mediile simple se calculează pentru serii cu frecvențe egale, de forma $X: \{x_i, n_i\}, n_1 = n_2 = \dots = n_i, i = \overline{1, n}$, și pentru serii simple, de forma $\{x_i\}, i = \overline{1, n}$.
18. b) Mediile ponderate se calculează pentru serii cu frecvențe diferite, de forma $X: \{x_i, n_i\}, n_1 \neq n_2 \neq \dots \neq n_i, i = \overline{1, m}$.
19. b) Media aritmetică reprezintă valoarea pe care ar purta-o fiecare unitate statistică dacă distribuția ar fi omogenă.
20. c) Media aritmetică simplă se calculează după relația $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$.
21. a) și c) Media aritmetică ponderată se calculează după relațiile:

$$\bar{x} = \frac{1}{\sum_i n_i} \sum_i x_i n_i, i = \overline{1, m}; \bar{x} = \frac{\sum_i x_i f_i}{\sum_i f_i} = 1, i = \overline{1, m}.$$
22. b) și c) În cazul seriilor de intervale de variație, media aritmetică se calculează după relația $\bar{x} = \frac{1}{\sum_i n_i} \sum_i x'_i n_i, i = \overline{1, k}, x'_i = \frac{1}{2}(x_{i-1} + x_i)$,
 respectiv după relația $\bar{x} = \frac{\sum_i x'_i f_i}{\sum_i f_i} = 1, i = \overline{1, m}, x'_i = \frac{1}{2}(x_{i-1} + x_i)$.
23. a) b) și c) Media aritmetică este o mărime normală, o mărime internă și o mărime translativă.

24. a) Relațiile $\sum_i (x_i - \bar{x}) = 0$, $\sum_i (x_i - \bar{x})n_i = 0$ arată că media este o mărime normală.

25. a) Media sumei a două variabile aleatoare independente, X și Y , este egală cu suma mediilor celor două variabile, adică $\overline{x+y} = \bar{x} + \bar{y}$.

26. b) Media produsului a două variabile aleatoare independente, X și Y , este egală cu produsul mediilor celor două variabile, adică $\overline{x \cdot y} = \bar{x} \cdot \bar{y}$.

27. b) Relațiile $\bar{x} = \bar{x}' + a$, $\bar{x} = \bar{x}'' \cdot k$, $\bar{x} = \frac{\sum_i x_i \frac{n_i}{c}}{\sum_i \frac{n_i}{c}}$, unde: $\bar{x}' = \frac{1}{n} \sum_i (x_i \pm a)n_i$,

iar $\bar{x}'' = \frac{1}{n} \sum_i \frac{x_i}{k} n_i$, arată că media este o mărime translativă.

28. a) și c) Calculul simplificat al mediei se poate efectua după relația

$$\bar{x} = \frac{\sum_i \left(\frac{x_i - a}{k} \right)}{n} \cdot k + a, \quad i = \overline{1, n}, \quad \text{respectiv} \quad \bar{x} = \frac{\sum_i \left(\frac{x_i - a}{k} \right) \cdot \frac{n_i}{c}}{\sum_i \frac{n_i}{c}} \cdot k + a,$$

$i = \overline{1, n}$.

29. b) Media unei caracteristici alternative se calculează după relația $p = \frac{n_i}{n}$.

30. b) și c) Media geometrică se calculează după relațiile: $\bar{x}_g = \sqrt[n]{\prod_{i=1}^n x_i}$, $n > 2$,

respectiv $\bar{x}_g = \sqrt[n_i]{\prod_{i=1}^{n_i} x_i^{n_i}}$.

31. b) Media geometrică a unei variabile Z , definită ca produs a două variabile, X și Y , este egală cu produsul mediilor geometrice ale celor două variabile, $\bar{Z}_g = \bar{X}_g \cdot \bar{Y}_g$.

32. a) Media geometrică a unei variabile Z , definită ca raport al două variabile, X și Y , este egală cu raportul mediilor geometrice ale celor două variabile, adică $\bar{Z} = \bar{X}_g / \bar{Y}_g$.

33. b) Modul reprezintă valoarea caracteristicii cel mai frecvent observate într-o distribuție.
34. c) Mediana reprezintă valoarea caracteristicii care împarte volumul colectivității în două părți egale.
35. a) Unitatea mediană corespunde locului medianei în distribuție.
36. a) și b) Pentru o serie statistică, între mărimile medii de calcul există următoarele relații: $\bar{x}_h < \bar{x}_g < \bar{x} < \bar{x}_p$, respectiv $\bar{x}_p > \bar{x} > \bar{x}_g > \bar{x}_h$.
37. a) și c) Mediana este o mărime medie de poziție și o mărime medie fundamentală.
38. a) și b) Modul se utilizează pentru caracterizarea tendinței centrale și a gradului de asimetrie.
39. b) și d) Quantilele sunt mărimi de poziție care împart colectivitatea în 4 părți egale, respectiv în n părți egale.
40. a) și c) Mediana este o mărime medie de poziție care împarte suma valorilor globale ale seriei în două părți egale.
41. a) Determinarea modului, în cazul seriilor de intervale inegale, presupune calculul frecvențelor reduse.

42. a) și c) Modul se află după relația de calcul $Mo = x_{i-1} + d \cdot \frac{\Delta_1}{\Delta_1 + \Delta_2}$,

respectiv după relația $Mo = x_{i-1} + d \cdot \frac{(n_i - n_{i-1})}{(n_i - n_{i-1}) + (n_i - n_{i+1})}$.

43. a) și b) Mediana se află după relația $Me = x_{i-1} + d \cdot \frac{\sum n_i - N_{i-1}}{n_i}$, respectiv

$$Me = x_{i-1} + d \cdot \frac{U^{Me} - N_{i-1}}{n_i}.$$

5.11. Recunoașteți exemplul corect

1. b) Rata medie a inflației $= \sqrt[3]{1,1 \cdot 1,1 \cdot 1,7} = 1,2718$ este o medie geometrică.
2. a) Salariul mediu pentru angajații din întreprinderea considerată s-a calculat ca o medie aritmetică.

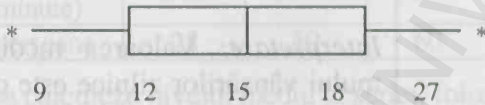
- 3. a) 25% din salariații unei întreprinderi au un salariu de cel mult 4 milioane lei. Această valoare reprezintă quartila unu.
- 4. c) 75% din salariații unei întreprinderi au un salariu de cel mult 7 milioane lei. Această valoare reprezintă quartila trei.
- 5. b) 50% din salariații unei întreprinderi au un salariu de cel mult 5 milioane lei. Această valoare reprezintă mediana.
- 6. b) și c) În distribuția următoare:

(x_{i-1}, x_i)	0-5	5-10	10-15	15-20
n_i	3	7	15	20

*Date convenționale

intervalul (15-20) reprezintă intervalul modal (dominant).

- 7. b) și c) Box-and-whiskers plot de mai jos arată o valoare mediană egală cu 15 și o valoare a quartilei trei egală cu 18.





5.13. Probleme

1. Valoarea vânzărilor înregistrate la firma „A”, pe parcursul a șapte zile, se prezintă astfel (în milioane lei): 200; 250; 750; 1001; 1520; 1800; 2300. Să se calculeze valoarea medie zilnică a vânzărilor realizate de firma „A”.

Valoarea medie a vânzărilor pe o zi se află calculând media aritmetică. Media aritmetică, în cazul unei serii simple, se calculează după relația:

$$\bar{x} = \frac{\sum_i x_i}{n}$$

Calcule:

$$\bar{x} = \frac{200+250+750+1001+1520+1800+2300}{7} = 1117,28 \text{ mil. lei}$$

Interpretare: Valoarea medie a volumului vânzărilor zilnice este de 1117,28 milioane lei.

2. Distribuția după volumul vânzărilor zilnice (milioane lei) a unităților comerciale dintr-un județ, în luna ianuarie 2004, se prezintă astfel:

Volumul vânzărilor	0-10	10-20	20-30	30-40
Numărul unităților comerciale	50	20	10	5

*Date convenționale

Să se calculeze nivelul mediu al vânzărilor zilnice.

Nivelul mediu al vânzărilor zilnice se află calculând media aritmetică. Media aritmetică, în cazul unei serii grupate pe intervale de variație, se calculează după relația:

$$\bar{x} = \frac{\sum_i x_i \cdot n_i}{\sum_i n_i}$$

Elementele de calcul sunt prezentate în tabelul 5.2.

$$\bar{x} = \frac{975}{85} = 11,47 \text{ mil. lei}$$

unde: $x'_i = \frac{x_{i-1} + x_i}{2}$

Interpretare: Valoarea medie a volumului vânzărilor zilnice este de 11,47 milioane lei.

Tabelul 5.2 Elemente de calcul

x_{i-1}, x_i	n_i	x'_i	$x'_i n_i$
0-10	50	5	250
10-20	20	15	300
20-30	10	25	250
30-40	5	35	175
Total	85	-	975

3. Cu privire la timpul de deplasare zilnică s-au înregistrat următoarele date pentru un eșanțion de 200 persoane:

Timpul de deplasare zilnică (în minute)	0-30	30-60	60-90	90-120	120-150	150 și peste
Număr de persoane	25	50	60	45	15	5

Se cere să se calculeze nivelul mediu al seriei, folosind calculul simplificat.

Calculul simplificat al mediei aritmetice:

$$\bar{x} = \frac{\sum \frac{x'_i - a}{k} \cdot \frac{n_i}{c} \cdot k + a}{\sum \frac{n_i}{c}}$$

unde:

$$x'_i = \frac{x_{i-1} + x_i}{2};$$

a reprezintă, de regulă, o valoare centrală a șirului ordonat al caracteristicii;
 k este o valoare care dă o simplificare maximă elementelor din coloana $x'_i - a$;
 c reprezintă divizorul comun al șirului frecvențelor.

Pentru această distribuție, considerăm: $a=75, k=30, c=5$.

Elementele de calcul sunt prezentate în tabelul 5.3.

Media aritmetică, folosind formula de calcul simplificat, este:

$$\bar{x} = -\frac{2}{40} \cdot 30 + 75 = 73,5;$$

$\bar{x} = 73,5$ minute

Tabelul 5.3. Elemente de calcul

$x_{i-1} - x_i$	n_i	x'_i	$x'_i - a$	$\frac{x'_i - a}{k}$	$\frac{n_i}{c}$	$\frac{x'_i - a}{k} \cdot \frac{n_i}{c}$
1	2	3	4	5	6	7
0 - 30	25	15	-60	-2	5	-10
30 - 60	50	45	-30	-1	10	-10
60 - 90	60	75	0	0	12	0
90 - 120	45	105	30	1	9	9
120 - 150	15	135	60	2	3	6
150 -	5	165	90	3	1	3
Total	200	-	-	-	40	-2

4. În urma controlului de calitate, într-un lot de 500 de piese se găsesc $n_1 = 30$ piese rebut. Să se calculeze procentul mediu al pieselor rebut din lot.

Nivelul mediu în cazul unei variabile alternative se calculează după relația:

$$p = \frac{n_1}{n} \cdot 100$$

Calcule:

$$p = \frac{30}{500} \cdot 100 = 6\%$$

Interpretare: Rezultatul obținut arată că în medie, la fiecare 100 de piese observate, 6 piese sunt rebut.

5. Cu privire la timpul de deplasare zilnică, s-au înregistrat următoarele date pentru un eșantion de 200 persoane:

Timpul de deplasare zilnică (în minute)	0-30	30-60	60-90	90-120	120-150	150 și peste
Număr de persoane	25	50	60	45	15	5

Se cere să se calculeze nivelul mediu al seriei, folosind media armonică, media geometrică, media pătratică și să se verifice relațiile dintre aceste medii.

Relații de calcul:
- media aritmetică:

$$\bar{x} = \frac{\sum_{i=1}^m x'_i n_i}{\sum_{i=1}^m n_i}$$

Elementele de calcul sunt prezentate în tabelul 5.4.

$$\bar{x} = \frac{14700}{200} = 73,5 \text{ minute}$$

- media armonică:

$$\bar{x}_h = \frac{\sum n_i}{\sum \frac{1}{x_i} n_i}$$

- media geometrică:

$$\bar{x}_g = \sqrt[m]{\prod_{i=1}^m x_i^{n_i}}$$

Când $\sum n_i > 3$, pentru a facilita calculul, se folosește expresia logaritmică:

$$\lg \bar{x}_g = \frac{1}{\sum n_i} \sum_{i=1}^m n_i \lg x_i$$

- media pătratică:

$$\bar{x}_p = \sqrt{\frac{\sum_{i=1}^m x_i^2 n_i}{\sum_{i=1}^m n_i}}$$

Pentru aceeași serie, între mărimile medii există următoarele relații:

$$\bar{x}_h < \bar{x}_g < \bar{x} < \bar{x}_p$$

$$\bar{x}_h = \frac{200}{4,15} = 48,19 \text{ minute}$$

În cazul nostru $\sum n_i = 200$, prin urmare:

$$\lg \bar{x}_g = \frac{1}{200} \cdot 358,53 = 1,7926$$

$$\bar{x}_g = \text{antilog}(\log \bar{x}_g) = \text{antilog} 1,7926 = 62,09;$$

$$\bar{x}_g = 62,09 \text{ minute.}$$

$$\bar{x}_p = \sqrt{\frac{134987}{200}} = 82,15;$$

$$\bar{x}_p = 82,15 \text{ minute.}$$

Între valorile mediilor calculate, pentru exemplul considerat, există relațiile:

$$48,19 < 62,09 < 73,5 < 82,15.$$

Tabelul 5.4. Elemente de calcul

$x_{i-1} - x_i$	n_i	x'_i	$x'_i n_i$	$\frac{n_i}{x'_i}$	$\log x'_i$	$n_i \log x'_i$	$(x'_i)^2$	$(x'_i)^2 n_i$
1	2	3	4	5	6	7	8	9
0- 30	25	15	375	1,67	1,176	29,40	225	5625
30- 60	50	45	2250	1,11	1,653	82,65	2025	101125
60- 90	60	75	4500	0,80	1,875	112,50	5625	337500
90-120	45	105	4725	0,43	2,021	90,94	11025	496125
125-150	15	135	2025	0,11	2,130	31,95	18225	273375
150 și peste	5	165	825	0,03	2,217	11,09	27225	136125
Total	200	-	14700	4,15	-	358,53	-	1349875

6. Distribuția familiilor dintr-un bloc după numărul de copii se prezintă astfel:

Număr copii	0	1	2	3	4	5
Număr familii	4	8	15	3	2	1

*Date convenționale

Să se calculeze valoarea modului.

Modul este valoarea caracteristicii corespunzătoare nivelului maxim ($n_i=n_{max}$) al efectivului.

Pentru distribuția considerată,
 $n_{max} = 15$.
Valoarea caracteristicii corespunzătoare efectivului maxim, $n_{max}=15$, este $x_i=2$ copii.
 $Mo=2$ copii
Interpretare: La nivelul eșantionului considerat, cele mai multe familii au 2 copii.

7. Distribuția după volumul vânzărilor zilnice (milioane lei) a unităților comerciale dintr-un județ, în luna ianuarie 2004, se prezintă astfel:

Volumul vânzărilor	0-10	10-20	20-30	30-40
Numărul unităților comerciale	20	50	10	5

*Date convenționale

Să se calculeze valoarea modului.

Într-o serie de intervale, aflarea modului presupune:
- depistarea valorii n_{max} ;
- aflarea intervalului modal, (x_{i-1}, x_i) ;
- interpolarea în intervalul modal după relația:

Valoarea efectivului maxim este $n_i=50$.
Intervalul modal corespunzător este (10; 20).

$$Mo = x_{i-1} + d \cdot \frac{\Delta_1}{\Delta_1 + \Delta_2},$$

unde:

$$\Delta_1 = n_i - n_{i-1}$$

$$\Delta_2 = n_i - n_{i+1}$$

$$Mo = 10 + 10 \cdot \frac{(50 - 20)}{(50 - 20) + (50 - 10)} =$$

$$= 14,28 \text{ mil. lei}$$

Interpretare: Volumul vânzărilor zilnice realizate de cele mai multe unități comerciale din localitatea considerată este de 14,28 milioane lei.

8. O firmă înregistrează, într-o perioadă de cinci zile, următoarele date privind volumul vânzărilor (milioane lei): 5, 7, 20, 14, 9. Să se afle valoarea medianei.

Pentru o serie simplă cu număr impar de termeni, ordonați în sens crescător, mediana se află depistând termenul central al seriei, în dreptul

U^{Me} :

$$U^{Me} = \frac{n+1}{2}.$$

Unitatea mediană este:

$$U^{Me} = (5+1)/2 = 3$$

Prin urmare, locul medianei este al treilea termen al șirului:

(5, 7, 9, 14, 20), deci $Me = 9$ milioane lei.

9. O firmă înregistrează, într-o perioadă de șase zile, următoarele date privind volumul vânzărilor (milioane lei): 10, 40, 20, 15, 22, 35. Să se afle valoarea medianei.

Pentru o serie simplă cu număr par de termeni, ordonați în sens crescător, aflarea medianei constă în calculul mediei aritmetice simple a celor doi termeni centrali ai seriei.

Șirul de date ordonate crescător este (10, 15, 20, 22, 35, 40).

$$Me = \frac{20 + 22}{2} = 21 \text{ milioane lei}$$

10. Distribuția familiilor dintr-un bloc după numărul de copii se prezintă astfel:

Număr copii	0	1	2	3	4	5
Număr familii	4	8	15	4	2	1

*Date convenționale

Să se calculeze valoarea mediane.

Pentru o serie cu $n_1 \neq n_2 \neq \dots \neq n_i$, aflarea mediane presupune următoarele operații:

- se calculează *efectivele cumulate*:

$$N_i = \sum_{h=1}^i n_h$$

- se calculează *unitatea mediană*:

$$U^{Me} = \frac{\sum_i n_i + 1}{2}$$

- se află *efectivul cumulat care verifică relația*: $N_i \geq U^{Me}$

- se află *valoarea caracteristicii corespunzătoare efectivului cumulat care verifică relația*: $N_i \geq U^{Me}$

Efectivele cumulate (N_i) sunt prezentate în tabelul de mai jos:

x_i	n_i	N_i
0	4	4
1	8	12
2	15	27
3	4	31
4	2	33
5	1	34
Total	34	-

$$U^{Me} = \frac{34+1}{2} = 17,5$$

$$(N_i=27) \geq (U^{Me}=17,5)$$

Prin urmare, $Me=2$ copii.

Interpretare: La nivelul eșantionului considerat, jumătate din familii au până la 2 copii, iar jumătate au peste 2 copii.

11. Distribuția muncitorilor unei firme după timpul consumat pentru realizarea produsului „A” se prezintă astfel:

Timpul necesar (min)	-30	30-60	60-90	90-120	120-150	150-
Număr muncitori	25	50	60	45	15	5

*Date convenționale

Să se calculeze valoarea mediane.

Pentru o serie de intervale cu $n_1 \neq n_2 \neq \dots \neq n_i$, aflarea medianeî presupune următoarele operații:

- se calculează *efectivele cumulate*:

$$N_i = \sum_{h=1}^i n_h$$

- se calculează *unitatea mediană*:

$$U^{Me} = \frac{\sum_i n_i}{2}$$

- se află *efectivul cumulat care verifică relația*: $N_i \geq U^{Me}$

- se află *intervalul median* corespunzător efectivului cumulat care verifică relația: $N_i \geq U^{Me}$

- se calculează *mediana* prin interpolare în intervalul median, după relația:

$$Me = x_{i-1} + d \cdot \frac{U^{Me} - N_{i-1}}{n_i}$$

Efectivele cumulate (N_i) sunt prezentate în tabelul de mai jos:

x_{i-1}, x_i	n_i	x_i	N_i
0-30	25	15	25
30-60	50	45	75
60-90	60	75	135
90-120	45	105	180
120-150	15	135	195
150-180	5	165	200
Total	200	-	-

$$U^{Me} = \frac{200}{2} = 100$$

$$(N_i = 135) \geq (U^{Me} = 100).$$

Prin urmare, *intervalul median* este intervalul (60 – 90).

$$Me = 60 + 30 \cdot \frac{100 - 75}{60} = 72,5 \text{ min.}$$

Interpretare: Jumătate din numărul total al muncitorilor au consumat pentru realizarea produsului „A” până la 72,5 minute, iar jumătate au consumat peste 72,5 minute.

12. Distribuția muncitorilor firmei „A” după timpul consumat pentru realizarea produsului „x” se prezintă astfel:

Timpul consumat (min)	-30	30-60	60-90	90-120	120-150	150-
Număr muncitori	25	50	60	45	15	5

*Date convenționale

Să se calculeze valoarea quartilei trei.

Pentru o serie de intervale cu $n_1 \neq n_2 \neq \dots \neq n_i$, aflarea quartilei

trei presupune următoarele operații:

- se calculează *efectivele cumulate*:

$$N_i = \sum_{h=1}^i n_h$$

- se calculează *unitatea quartilică*

$$\text{trei: } U^{Q_3} = \frac{3 \cdot \sum n_i}{4}$$

- se află *efectivul cumulat care verifică relația*: $N_i \geq U^{Q_3}$

- se află *intervalul quartilic trei* corespunzător efectivului cumulat care verifică relația: $N_i \geq U^{Q_3}$

- se calculează *quartila trei* prin *interpolare* în intervalul quartilic trei, după relația:

$$Q_3 = x_{i-1} + d \cdot \frac{U^{Q_3} - N_{i-1}}{n_{Q_3}}$$

Efectivele cumulate (N_i) sunt prezentate în tabelul de mai jos:

x_{i-1}, x_i	n_i	x_i	$N_i(\downarrow)$
0-30	25	15	25
30-60	50	45	75
60-90	60	75	135
90-120	45	105	180
120-150	15	135	195
150-180	5	165	200
Total	200	-	-

$$U^{Q_3} = \frac{3 \cdot 200}{4} = 150$$

$$(N_i = 180) \geq (U^{Q_3} = 150)$$

Prin urmare, *intervalul quartilic 3* este intervalul (90-120).

$$Q_3 = 90 + 30 \cdot \frac{150 - 135}{45} = 100 \text{ min.}$$

Interpretare: 75% din numărul total al muncitorilor consumă cel mult 100 minute pentru realizarea unei piese.

13. Distribuția muncitorilor unei firme după timpul consumat pentru realizarea produsului „A” se prezintă astfel:

Timpul necesar (min)	-30	30-60	60-90	90-120	120-150	150-
Număr muncitori	25	50	60	45	15	5

*Date convenționale

Să se calculeze valoarea quartilei unu.

Pentru o serie de intervale cu $n_1 \neq n_2 \neq \dots \neq n_i$, aflarea quartilei unu presupune următoarele operații:

Efectivele cumulate (N_i) sunt prezentate în tabelul de mai jos:

- se calculează *efectivele cumulate*:

$$N_i = \sum_{h=1}^i n_h$$

- se calculează *unitatea quartilică*

unu:
$$U^{Q_1} = \frac{1 \cdot \sum n_i}{4}$$

- se află *efectivul cumulat care verifică relația*: $N_i \geq U^{Q_1}$

- se află *intervalul quartilic unu corespunzător efectivului cumulat care verifică relația*: $N_i \geq U^{Q_1}$

- se calculează *quartila unu prin interpolare* în intervalul quartilic unu, după relația:

$$Q_1 = x_{i-1} + d \cdot \frac{U^{Q_1} - N_{i-1}}{n_{Q_1}}$$

x_{i-1}, x_i	n_i	x_i	N_i
0-30	25	15	25
30-60	50	45	75
60-90	60	75	135
90-120	45	105	180
120-150	15	135	195
150-180	5	165	200
Total	200	-	-

$$U^{Q_1} = \frac{1 \cdot 200}{4} = 50$$

$$(N_i = 75) \geq (U^{Q_1} = 50)$$

Prin urmare, *intervalul quartilic unu* este intervalul (30-60).

$$Q_1 = 30 + 30 \cdot \frac{50 - 25}{50} = 45 \text{ min.}$$

Interpretare: 25% din numărul total al muncitorilor consumă cel mult 45 minute pentru realizarea unei piese.

14. Distribuția muncitorilor unei firme după timpul consumat pentru realizarea produsului „A“ se prezintă astfel:

Timpul necesar (min)	-30	30-60	60-90	90-120	120-150	150 -
Număr muncitori	25	50	60	45	15	5

**Date convenționale*

Să se calculeze valoarea decilei unu.

Pentru o serie de intervale cu $n_1 \neq n_2 \neq \dots \neq n_i$, aflarea decilei unu presupune următoarele operații:
- se calculează *efectivele cumulate*:

Efectivele cumulate (N_i) sunt prezentate în tabelul de mai jos:

$$N_i = \sum_{h=1}^i n_h$$

- se calculează *unitatea decilică*

$$\text{unu: } U^{D1} = \frac{1 \cdot \sum_i n_i}{10}$$

- se află *efectivul cumulat care verifică relația: $N_i \geq U^{D1}$*

- se află *intervalul decilic unu* corespunzător efectivului cumulat care verifică relația: $N_i \geq U^{D1}$

- se calculează *valoarea decilei unu* prin *interpolare* în intervalul decilic unu, după relația:

$$D_1 = x_{i-1} + d \cdot \frac{U^{D1} - N_{i-1}}{n_{D1}}$$

x_{i-1}, x_i	n_i	x_i	N_i
0-30	25	15	25
30-60	50	45	75
60-90	60	75	135
90-120	45	105	180
120-150	15	135	195
150-180	5	165	200
Total	200	-	-

$$U^{D1} = \frac{1 \cdot 200}{10} = 20$$

$$(N_i = 25) \geq (U^{D1} = 20)$$

Prin urmare, *intervalul decilic unu* este intervalul (0-30).

$$D_1 = 0 + 30 \cdot \frac{20 - 0}{25} = 24 \text{ min.}$$

Interpretare: 10% din numărul total al muncitorilor consumă cel mult 24 minute pentru realizarea unei piese.

15. Distribuția muncitorilor unei firme după timpul consumat pentru realizarea produsului „A” se prezintă astfel:

Timpul necesar (min)	-30	30-60	60-90	90-120	120-150	150-
Număr muncitori	25	50	60	45	15	5

*Date convenționale

Să se calculeze valoarea decilei nouă.

Pentru o serie de intervale cu $n_1 \neq n_2 \neq \dots \neq n_i$, aflarea decilei nouă presupune următoarele operații:

- se calculează *efectivele cumulate*:

Efectivele cumulate (N_i) sunt prezentate în tabelul de mai jos:

$$N_i = \sum_{h=1}^i n_h$$

- se calculează *unitatea decilică*

nouă: $U^{D_9} = \frac{9 \cdot \sum_i n_i}{10}$

- se află *efectivul cumulat* care verifică relația: $N_i \geq U^{D_9}$

- se află *intervalul decilic nouă* corespunzător efectivului cumulat care verifică relația: $N_i \geq U^{D_9}$

- se calculează valoarea decilei nouă prin *interpolare* în intervalul decilic nouă, după relația:

$$D_9 = x_{i-1} + d \cdot \frac{U^{D_9} - N_{i-1}}{n_{D_9}}$$

x_{i-1}, x_i	n_i	x_i	N_i
0-30	25	15	25
30-60	50	45	75
60-90	60	75	135
90-120	45	105	180
120-150	15	135	195
150-180	5	165	200
Total	200	-	-

$$U^{D_9} = \frac{9 \cdot 200}{10} = 180$$

$$(N_i = 180) \geq (U^{D_9} = 20)$$

Prin urmare, *intervalul decilic 9* este intervalul (90-120).

$$D_9 = 90 + 30 \cdot \frac{180 - 135}{45} = 120 \text{ min.}$$

Interpretare: 90% din numărul total al muncitorilor consumă pentru realizarea unei piese cel mult 120 minute.

16. Distribuția muncitorilor unei firme după timpul consumat pentru realizarea produsului „A” se prezintă astfel:

Timpul necesar (min)	-30	30-60	60-90	90-120	120-150	150 -
Număr muncitori	25	50	60	45	15	5

*Date convenționale

Să se calculeze valoarea medialei.

Pentru o serie de intervale cu $n_1 \neq n_2 \neq \dots \neq n_i$, aflarea medialei presupune următoarele operații:

- se calculează *produsele* ($x_i n_i$);
- se calculează *produsele* ($x_i n_i$)

Produsele ($x_i n_i$) cumulate sunt prezentate în tabelul de mai jos:

cumulate: $L_h = \sum_{i=1}^h x_i' n_i$;

- se calculează *unitatea medială*

$$U^{Ml} = \frac{\sum x_i' n_i}{2}$$

- se află *valoarea* (L_h) care verifică relația: $L_h \geq U^{Ml}$

- se află *intervalul medial* corespunzător valorii: $L_h \geq U^{Ml}$

- se calculează *valoarea medialei* prin *interpolare* în intervalul medial, după relația:

$$Ml = x_{i-1} + d \cdot \frac{U^{Ml} - L_{i-1}}{x_i n_i - L_{i-1}}$$

x_{i-1}, x_i	n_i	x_i	$x_i' n_i$	L_h
0-30	25	15	375	375
30-60	50	45	2250	2625
60-90	60	75	4500	7125
90-120	45	105	4725	11850
120-150	15	135	2025	13875
150-180	5	165	825	14700
Total	200	-	14700	-

$$U^{Ml} = \frac{14700}{2} = 7350$$

$$(L_h = 11850) \geq (U^{Ml} = 7350)$$

Prin urmare, *intervalul medial* este intervalul (90-120).

$$Ml = 90 + 30 \cdot \frac{7350 - 7125}{4725} = 91,43 \text{ min.}$$

Interval (min)	Frecvență	Frecvență cumulat
0-30	25	25
30-60	50	75
60-90	60	135
90-120	45	180
120-150	15	195
150-180	5	200

*Date complementare

*Date complementare

Capitolul 6

INDICATORI AI DISPERSIEI, ASIMETRIEI ȘI BOLTIRII

- ◆ Indicatori simpli și sintetici ai dispersiei
- ◆ Proprietăți și calculul simplificat al varianței
- ◆ Indicatori ai dispersiei unei variabile alternative
- ◆ Măsurarea dispersiei în sistemul medianei
- ◆ Dispersia unei variabile atributive (categoriale)
- ◆ Indicatori ai dispersiei unei distribuții bidimensionale
- ◆ Indicatori ai formei
- ◆ Calculul indicatorilor dispersiei, asimetriei și boltirii în SPSS
- ◆ Alegeți răspunsul corect
- ◆ Recunoașteți exemplul corect
- ◆ Răspunsuri corecte
- ◆ Probleme

6.1. Indicatori simpli și sintetici ai dispersiei

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 148-155)

- **Indicatorii simpli ai dispersiei**

- Amplitudinea variației
- Abaterea individuală

- **Indicatori sintetici ai dispersiei**

- Abaterea medie liniară
- Varianța
- Abaterea medie pătratică (deviația standard)
- Intervale de variație (Figura 6.1)
- Coeficientul de variație

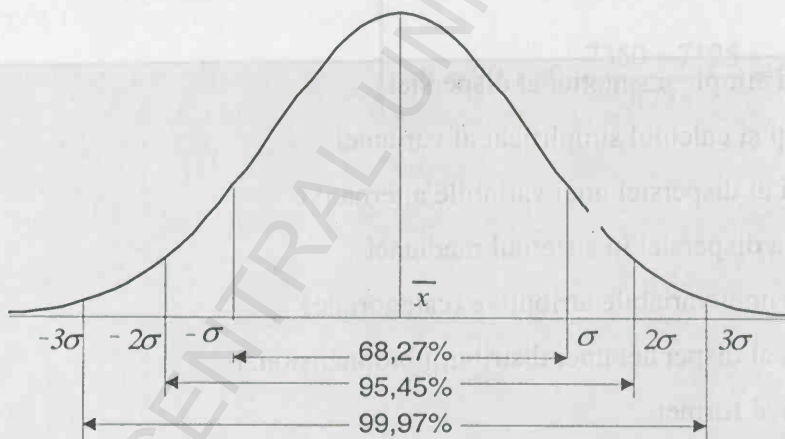


Fig. 6.1 Intervale de variație

6.2. Proprietăți și calculul simplificat al varianței

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 155-159)

- **Calculul simplificat al varianței**

- pentru o serie simplă

$$\sigma^2 = \frac{\sum_i \left(\frac{x_i - a}{k} \right)^2}{n} \cdot k^2 - (\bar{x} - a)^2$$

- pentru o serie cu frecvențe diferite

$$\sigma^2 = \frac{\sum_i \left(\frac{x_i - a}{k} \right)^2 \cdot \frac{n_i}{c}}{\sum_i \frac{n_i}{c}} \cdot k^2 - (\bar{x} - a)^2$$

6.3. Indicatori ai dispersiei unei variabile alternative

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 160-161)

Varianța unei variabile alternative este:

$$\sigma_p^2 = p \cdot q = p(1 - p)$$

6.4. Măsurarea dispersiei în sistemul medianei

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 161-165)

- Intervalul interquartilic (Vezi Figura 6.2)
- Semiinterquartila
- Coeficientul de variație interquartilic
- Intervalul interdecilic
- Coeficientul de variație interdecilic



Fig. 6.4. Asimetrie la dreapta



Fig. 6.5. Asimetrie la stânga

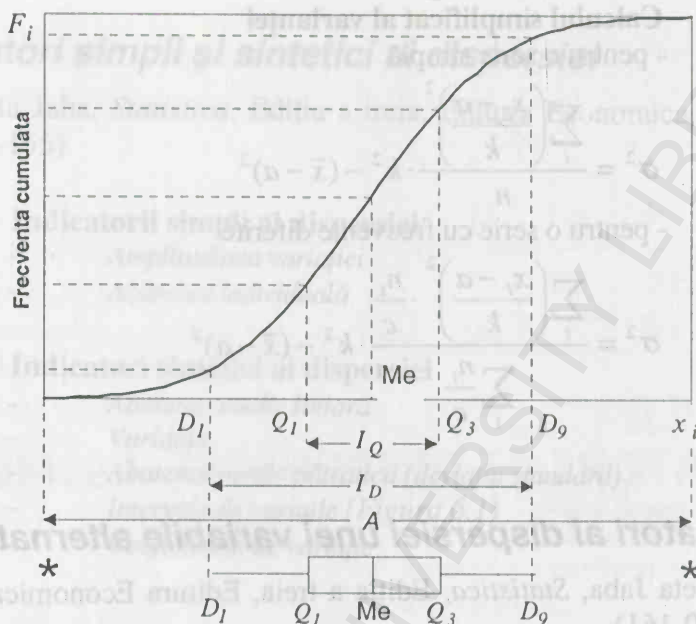


Fig. 6.2. Intervale interquantilice

6.5. Indicatori factoriali ai dispersiei

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 166-176)

- **Indicatorii factoriali ai dispersiei**

- varianța de grupă (σ_j^2), sau varianța intragrupă (varianța condiționată)
- media varianțelor de grupă ($\overline{\sigma^2}$)
- varianța mediilor de grupă față de media generală (δ^2), numită și varianța inter-grupe (varianța mediilor condiționate)
- varianța generală, σ^2 , (varianța marginală)

- **Covarianța**

6.6. Indicatori ai formei

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 176-183)

Asimetria exprimă modul de repartizare a frecvențelor de o parte și alta a valorii centrale a unei serii.

- **Aprecierea grafică a asimetriei**
 - curba frecvențelor
 - box-and-whiskers plot

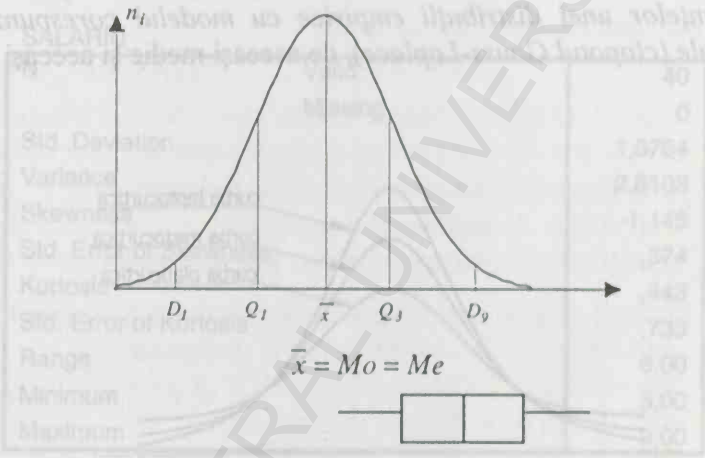


Fig. 6.3. Distribuție simetrică

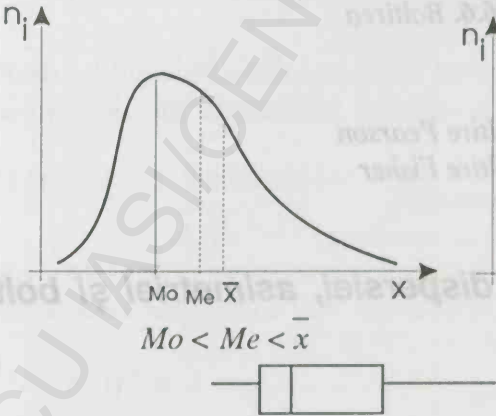


Fig. 6.4. Asimetrie la dreapta

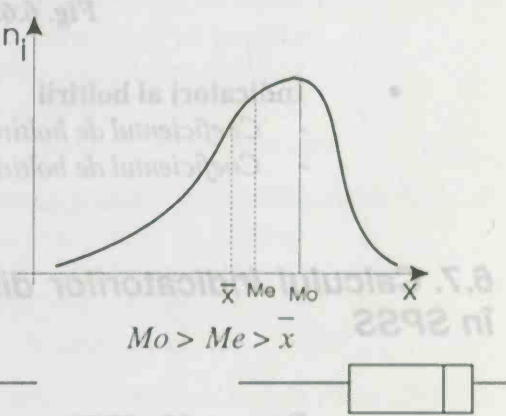


Fig. 6.5. Asimetrie la stânga

- **Indicatori ai asimetriei**
 - Asimetrie în mărime absolută
 - Coeficientul Yule
 - Coeficientul empiric de asimetrie Pearson
 - Coeficientul de asimetrie Pearson
 - Coeficientul de asimetrie Fisher

Boltirea (aplatizarea) exprimă măsura aglomerării frecvențelor în zona centrală.

- **Aprecierea grafică a boltirii** presupune compararea curbei frecvențelor unei distribuții empirice cu modelul corespunzător legii normale (clopotul Gauss-Laplace), de aceeași medie și aceeași dispersie.

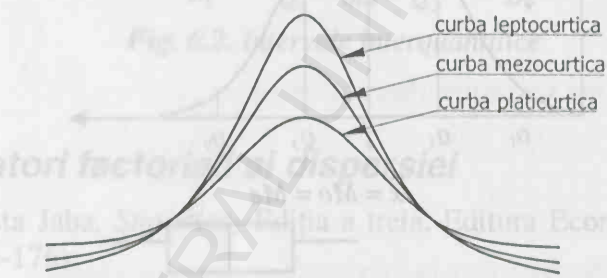


Fig. 6.6. Boltirea

- **Indicatori ai boltirii**
 - Coeficientul de boltire Pearson
 - Coeficientul de boltire Fisher

6.7. Calculul indicatorilor dispersiei, asimetriei și boltirii în SPSS

- **Demersul în SPSS**
 - Meniul *Analyse*, opțiunea *Descriptives Statistics*, comanda *Frequencies*

- În fereastra *Frequencies*, se alege *Statistics: Dispersion* (*Variance, Standard Deviation, Minimum, Maximum*) și *Distribution* (*Skewness, Kurtosis*).

- **Rezultate în SPSS**

- *Rezultatele în SPSS (output-ul „Statistics”) prezintă valorile indicatorilor dispersiei (Standard Deviation, Variance, Minimum, Maximum), asimetriei (Skewness) și boltirii (Kurtosis).*

Statistics		
SALARIU		
N	Valid	40
	Missing	0
Std. Deviation		1,6764
Variance		2,8103
Skewness		-1,145
Std. Error of Skewness		,374
Kurtosis		,443
Std. Error of Kurtosis		,733
Range		6,00
Minimum		3,00
Maximum		9,00



6.8. Alegeți răspunsul corect

1. Fenomenul de dispersie exprimă
 - a) gradul de împrăștiere a valorilor individuale ale unei variabile în jurul valorii medii
 - b) împrăștierea tuturor nivelurilor individuale ale unei variabile față de nivelul maxim al variabilei
 - c) gradul de împrăștiere a unităților statistice în jurul valorii medii
2. Dispersia este datorată acțiunii factorilor
 - a) aleatori
 - b) esențiali
 - c) centrali
3. Dispersia, în cazul seriilor unidimensionale, este măsurată cu ajutorul
 - a) unui sistem de indicatori simpli ai dispersiei
 - b) unui sistem de indicatori factoriali ai dispersiei
 - c) unui sistem de indicatori ai dinamicii
 - d) unui sistem de indicatori sintetici ai dispersiei
4. Dispersia, în cazul seriilor multivariate, este măsurată cu ajutorul
 - a) unui sistem de indicatori simpli ai dispersiei
 - b) unui sistem de indicatori factoriali ai dispersiei
 - c) unui sistem de indicatori ai dinamicii
 - d) unui sistem de indicatori sintetici ai dispersiei
5. Indicatorii simpli ai dispersiei măsoară
 - a) câmpul de împrăștiere al variabilei
 - b) împrăștierea fiecărui nivel individual față de nivelul lor mediu
 - c) împrăștierea tuturor nivelurilor individuale ale unei variabile față de nivelul lor mediu
6. Indicatorii sintetici ai dispersiei măsoară
 - a) câmpul de împrăștiere al variabilei
 - b) împrăștierea fiecărui nivel individual față de nivelul lor mediu
 - c) împrăștierea tuturor nivelurilor individuale ale unei variabile față de nivelul lor mediu
7. Indicatorii simpli ai dispersiei sunt
 - a) amplitudinea variației
 - b) abaterea individuală
 - c) coeficientul de variație

d) varianța

8. Amplitudinea variației se calculează după relația

a) $A = x_{\max} - x_{\min}$

b) $A_{\%} = \frac{x_{\max} - x_{\min}}{\bar{x}} \cdot 100$

c) $A_{\%} = \frac{\sigma}{\bar{x}} \cdot 100$

9. Abaterea individuală se calculează după relația

a) $d_i = x_i - \bar{x}$

b) $d_{i\%} = \frac{x_i - \bar{x}}{\bar{x}} \cdot 100$

c) $d_{i\%} = \frac{\sigma}{\bar{x}} \cdot 100$

10. Indicatorii sintetici ai dispersiei sunt

a) abaterea medie liniară și abaterea medie pătratică

b) varianța și coeficientul de variație

c) amplitudinea variației și abaterea individuală

11. Abaterea medie liniară se calculează după relațiile

a) $\bar{d} = \frac{\sum_i (x_i - \bar{x})}{n}$

b) $\bar{d} = \frac{\sum_i |x_i - \bar{x}| \cdot n_i}{n}$

c) $\bar{d} = \frac{\sum_i |x_i - \bar{x}|^2}{n}$

12. Varianța se calculează după relația

a) $\sigma^2 = \frac{\sum_i (x_i - \bar{x})^2}{n}$

b) $\sigma^2 = \frac{\sum_i (x_i - \bar{x})^2 \cdot n_i}{\sum_i n_i}$

$$c) \sigma^2 = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n}}$$

13. Abaterea medie pătratică se calculează după relațiile

$$a) \sigma = \sqrt{\frac{\sum_i d_i^2 \cdot n_i}{\sum_i n_i}}$$

$$b) \sigma = \sqrt{\sigma^2}$$

$$c) \sigma = \frac{\sum_i d_i^2 \cdot n_i}{\sum_i n_i}$$

14. Intervalul mediu de variație este definit de relațiile

$$a) \bar{x} \pm \bar{d}$$

$$b) \bar{x} \pm \sigma$$

$$c) \bar{x} \leq \sigma$$

15. Într-o distribuție normală standard, în intervalul mediu ($\bar{x} \pm \sigma$) apar

a) 68,27% din cazuri

b) 50% din cazuri

c) 90% din cazuri

16. Coeficientul de variație se calculează după relația

$$a) v = \frac{d_i}{\bar{x}} \cdot 100;$$

$$b) v = \frac{\bar{d}}{\bar{x}} \cdot 100;$$

$$c) v = \frac{\sigma}{\bar{x}} \cdot 100$$

17. Coeficientul de variație poate lua valori cuprinse între

$$a) 0 < v < 100\%$$

$$b) -3 < v < 3$$

$$c) -0.3 < v < 0.3$$

18. Varianța este egală cu

$$a) \sigma^2 = \bar{x}^2 - (\bar{x})^2$$

$$b) \sigma^2 = \bar{x}^2 + (\bar{x})^2$$

$$c) \sigma^2 = \frac{1}{\sum_i n_i} \cdot \sum_i (x_i - a)^2 \cdot n_i - (\bar{x} - a)^2$$

19. Varianța este egală cu

a) $\sigma^2 = \sigma_x^2 \cdot k^2$, unde $\sigma_x^2 = \frac{1}{n} \sum_i \left(\frac{x_i - \bar{x}}{k} \right)^2$

b) $\sigma_x^2 = \frac{1}{n} \sum_i \left(\frac{x_i - a}{k} \right)^2 \cdot k^2 - (\bar{x} - a)^2$

c) $\sigma_x^2 = \frac{1}{n} \sum_i \left(\frac{x_i - a}{k} \right)^2 \cdot k^2 + (\bar{x} - a)^2$

20. Varianța unei caracteristici alternative este definită de relația

a) $\sigma_p^2 = p \cdot q$

b) $\sigma_p^2 = p \cdot (1 - p)$

c) $\sigma_p^2 = \sqrt{p \cdot q}$

21. Indicatorii dispersiei în sistemul median sunt

a) abaterea interquartilică, semiinterquartilia, coeficientul de variație interquartilic

b) abaterea medie interdecilică, coeficientul de variație interdecilic

c) abaterea medie liniară, mediana, coeficientul de corelație

22. Abaterea medie interquartilică se calculează după relația

a) $A_Q = Q_3 - Q_1$

b) $A_Q = \frac{Q_3 - Q_1}{2}$

c) $A_Q = \left(\frac{\frac{Q_3 - Q_1}{2}}{Me} \right) \cdot 100$

23. Coeficientul de variație interquartilic se calculează după relația

a) $V_Q = Q_3 - Q_1$

b) $V_Q = \frac{Q_3 - Q_1}{2}$

$$c) V_Q = \left(\frac{\frac{Q_3 - Q_1}{2}}{Me} \right) \cdot 100$$

24. Indicatorii factoriali ai dispersiei sunt:

- a) varianța de grupă (σ_j^2)
- b) media varianțelor de grupă ($\overline{\sigma^2}$)
- c) varianța mediilor de grupă față de media generală (δ^2)
- d) varianța generală (σ^2)

25. Varianța de grupă măsoară

- a) variația sub influența factorilor aleatori (variația reziduală)
- b) variația sub influența factorilor esențiali
- c) variația sub influența factorilor aleatori și esențiali

26. Varianța mediilor de grupă față de media generală (varianța intergrupe) măsoară

- a) variația sub influența factorilor întâmplători (variația reziduală)
- b) variația sub influența factorilor esențiali
- c) variația sub influența factorilor întâmplători și esențiali

27. Varianța generală măsoară

- a) variația sub influența factorilor întâmplători (variația reziduală)
- b) variația sub influența factorilor esențiali
- c) variația sub influența factorilor întâmplători și esențiali

28. Sunt corecte relațiile

- a) $\sigma^2 = \overline{\sigma^2} + \delta^2$
- b) $\overline{\sigma^2} = \sigma^2 - \delta^2$
- c) $\sigma^2 = \overline{\sigma^2} - \delta^2$

29. Varianța intragrupă (media varianțelor de grupă) se calculează după relațiile

$$a) \overline{\sigma_p^2} = \frac{\sum_j p_j q_j n_{\cdot j}}{\sum_j n_{\cdot j}}, \quad j = \overline{1, k}$$

b) $\overline{\sigma^2} = \frac{\sum_j \sigma_j^2 n_{\cdot j}}{\sum_j n_{\cdot j}}$, unde $\sigma_j^2 = \frac{1}{n_{\cdot j}} \sum_i (x_i - \overline{x_j})^2 n_{ij}$

c) $\overline{\sigma^2} = \sigma^2 + \delta^2$

30. Variația intergrupe (varianța mediilor de grupă față de media generală) se calculează după relațiile

a) $\delta_p^2 = \frac{\sum_j (p_j - p)^2 n_{\cdot j}}{\sum_j n_{\cdot j}}$

b) $\delta^2 = \frac{\sum_j (\overline{x_j} - \overline{x})^2 n_{\cdot j}}{\sum_j n_{\cdot j}}$

c) $\delta^2 = \frac{\sum_i (x_i - \overline{x})^2 n_{i\cdot}}{n_{\cdot\cdot}}$

31. În cazul unei distribuții bidimensionale, cu ambele variabile exprimate cifric, se calculează

- a) media și varianța marginală
- b) medii și varianțe condiționate
- c) medii și varianțe specifice

32. Variația condiționată (varianța de grupă) se calculează după relația

a) $\sigma_j^2 = \frac{1}{n_{\cdot j}} \sum_i (x_i - \overline{x_j})^2 n_{ij}$, pentru variabila X, condiționată de $Y = y_j$

b) $\sigma_i^2 = \frac{1}{n_{i\cdot}} \sum_j (y_j - \overline{y_i})^2 n_{ij}$, pentru variabila Y, condiționată de $X = x_i$

c) $\sigma_j^2 = \frac{1}{n_{\cdot j}} \sum_i (x_i - \overline{x_j})^2 n_{ij}$, pentru variabila Y, condiționată de $X = x_i$

33. Variația generală se calculează după relația

a) $\sigma^2 = \frac{1}{n_{\cdot\cdot}} \sum_i (x_i - \overline{x})^2 n_{i\cdot}$, cu $i = \overline{1, m}$

b) $\sigma^2 = \frac{1}{n_{\cdot\cdot}} \sum_i x_i^2 n_{i\cdot} - (\overline{x})^2$, cu $i = \overline{1, m}$

$$c) \sigma^2 = \frac{1}{n_{..}} \sum \sigma_j^2 n_{.j} + \frac{1}{n_{..}} \sum (\bar{x}_j - \bar{x})^2 n_{.j}$$

34. Covarianța a două variabile, X și Y, este o măsură a

- a) covariației
- b) variației simultane a celor două variabile
- c) variației independente

35. Covarianța este definită de relația

$$a) \text{cov}(X, Y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n}, \quad i = \overline{1, n}$$

$$b) \text{cov}(X, Y) = \frac{1}{n} \sum_i \sum_j (x_i - \bar{x})(y_j - \bar{y}) \cdot n_{ij}, \quad i = \overline{1, k}, \quad j = \overline{1, p}$$

$$c) \text{cov}(X, Y) = \frac{1}{n_{..}} \sum (x_i - y_i) \cdot n_{ij}, \quad i = \overline{1, k}, \quad j = \overline{1, p}$$

36. Forma unei distribuții statistice unidimensionale poate fi apreciată cu ajutorul indicatorilor de

- a) asimetrie
- b) boltire
- c) intensitate

37. Indicatorii de asimetrie dau informații asupra modului de repartizare a frecvențelor de o parte sau alta a

- a) valorii centrale a unei serii
- b) valorii varianței
- c) valorii coeficientului de asimetrie

38. Indicatorii de boltire exprimă

- a) modul de repartizare a frecvențelor de o parte sau alta a valorii minime a unei serii
- b) măsura aglomerării frecvențelor în zona centrală, pe lângă medie
- c) măsura alungirii curbei frecvențelor spre plus infinit

39. Asimetria reprezintă

- a) deviația de la forma simetrică a distribuției
- b) deviația de la originea axelor de coordonate
- c) deviația de la valoarea centrală

40. Aprecierea grafică a asimetriei poate fi făcută cu ajutorul

- a) curbei de concentrare

- b) curbei frecvențelor
- c) diagramei box-and-whiskers

41. Într-o distribuție unidimensională simetrică, între cele trei mărimi medii fundamentale există relația

- a) $\bar{x} < Do < Me$, iar perechile de quantile Q_1 și Q_3 , respectiv D_1 și D_9 sunt echidistante față de valoarea centrală
- b) $\bar{x} = Do = Me$, iar perechile de quantile Q_1 și Q_3 , respectiv D_1 și D_9 sunt echidistante față de valoarea centrală
- c) $\bar{x} > Me > Do$, iar perechile de quantile Q_1 și Q_3 , respectiv D_1 și D_9 sunt echidistante față de valoarea centrală

42. Asimetria se calculează după relația

- a) $As = \bar{x} - Mo$
- b) $As = 3(\bar{x} - Me)$
- c) $As = 2(\bar{x} - Mo)$

43. Coeficientul de asimetrie Yule (Bowley) se calculează după relația

- a) $Cay = \frac{(q_2 - q_1)}{(q_2 + q_1)}$
- b) $Cay = \frac{(q_2 + q_1)}{(q_2 - q_1)}$
- c) $Cay = \frac{(q_3 - q_1)}{(q_3 + q_1)}$

44. Coeficientul empiric de asimetrie Pearson se calculează după relația

- a) $Cas = \frac{\bar{x} - Do}{\sigma}$
- b) $Cas = \frac{\mu_3^2}{\mu_2^3}$
- c) $Cas = \frac{\mu_3}{\sqrt{\mu_2^3}}$

45. Coeficientul de asimetrie Pearson, determinat pe baza momentelor centrate, se calculează după relația

- a) $\beta_1 = \frac{\bar{x} - Do}{\sigma}$
- b) $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$

$$c) \beta_1 = \frac{\mu_3}{\sqrt{\mu_2^3}}$$

46. O distribuție empirică, în comparație cu distribuția normală de aceeași medie și varianță, este leptocurtică atunci când prezintă

- a) o variație slabă a variabilei X însoțită de o variație puternică a frecvenței
- b) o variație puternică a variabilei X însoțită de o variație slabă a frecvenței
- c) o variație egală cu zero

47. O distribuție empirică, în comparație cu distribuția normală de aceeași medie și varianță, este platycurtică atunci când prezintă

- a) o variație slabă a variabilei X însoțită de o variație puternică a frecvenței
- b) o variație puternică a variabilei X însoțită de o variație slabă a frecvenței
- c) o variație egală cu zero

48. Coeficientul de boltire Pearson se calculează după relația

$$a) \beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$b) \beta_2 = \frac{\mu_3}{\mu_2^3}$$

$$c) \beta_2 = \frac{\mu_3}{\sigma^3}$$

49. Coeficientul de boltire Fisher se calculează după relația

$$a) \gamma_2 = \beta_2 - 3$$

$$b) \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$$

$$c) \gamma_2 = \frac{\mu_3}{\mu_2^3}$$

50. O distribuție este mezocurtică atunci când

- a) $\beta_2 = 3$, $\gamma_2 = 0$
- b) $\beta_2 > 3$, $\gamma_2 > 0$
- c) $\beta_2 < 3$, $\gamma_2 < 0$

51. O distribuție este leptocurtică atunci când

- a) $\beta_2 = 3$, $\gamma_2 = 0$
- b) $\beta_2 > 3$, $\gamma_2 > 0$
- c) $\beta_2 < 3$, $\gamma_2 < 0$

52. O distribuție este platycurtică atunci când

- a) $\beta_2 = 3$, $\gamma_2 = 0$
- b) $\beta_2 > 3$, $\gamma_2 > 0$
- c) $\beta_2 < 3$, $\gamma_2 < 0$

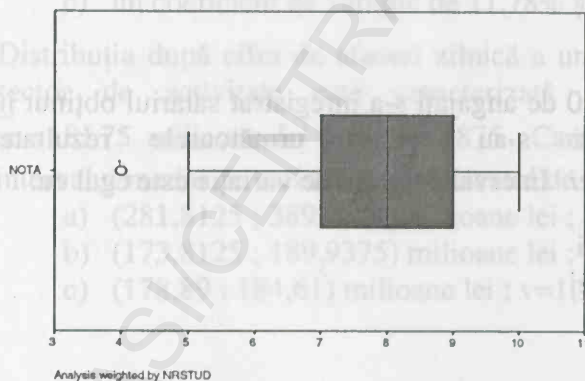


6.9. Recunoașteți exemplul corect

1. Distribuția unui eșantion de 200 de firme din județul Iași după profitul lunar, obținut în anul 2004, este caracterizată prin următoarele valori: $\bar{x} = 180$ milioane lei și $\sigma = 19$ milioane lei. Abaterea medie pătratică arată că
 - a) fiecare firmă obține un profit lunar care se abate de la profitul mediu cu 19 milioane lei, în sens pozitiv și negativ
 - b) fiecare firmă obține un profit lunar care se abate în medie, de la profitul mediu, cu 19 milioane lei, în sens pozitiv și negativ
 - c) fiecare firmă obține un profit lunar mai mare decât profitul mediu cu 19 milioane lei
2. Magazinele unei firme au practicat, pentru un produs, următoarele prețuri de vânzare: 100, 110, 120, 130, 140 mii lei. Pentru indicatorii dispersiei s-au obținut următoarele valori: $A\% = 33,3\%$; $\bar{d} = 12$ mii lei; $\sigma = 14.142$ mii lei, $v = 11,78\%$ și arată
 - a) un câmp de variație de 40 mii lei și o colectivitate omogenă
 - b) un câmp de variație de 80 mii lei și o colectivitate eterogenă
 - c) un coeficient de variație de 11,78% și o medie reprezentativă
3. Distribuția după cifra de afaceri zilnică a unui eșantion de firme dintr-un sector de activitate este caracterizată prin următoarele rezultate: $\bar{x} = 181,75$ milioane lei; $\sigma^2 = 8,1875$. Care sunt valorile corecte pentru intervalul mediu de variație și coeficientul de variație ?
 - a) (281,8125 ; 389,9375) milioane lei ; $v = 1,57\%$
 - b) (173,8125 ; 189,9375) milioane lei ; $v = 57\%$
 - c) (178,89 ; 184,61) milioane lei ; $v = 1,57\%$

4. Distribuția muncitorilor dintr-o secție de producție, după timpul consumat pentru realizarea unui produs, este caracterizată prin următoarele valori:
 $n=100$ muncitori, $\bar{x}=137,4$ minute, $\sum_i (x_i - \bar{x})^2 n_i = 16224$,
 $\sum_i x_i^2 n_i = 1.904.100$. Rezultatele corecte sunt
- $\sigma = 137,4$ minute; $v = 100\%$
 - $\sigma = 12,74$ minute; $v = 9,27\%$
 - $\sigma = 162,24$ minute; $v = 10\%$
5. Pentru o distribuție statistică după o variabilă X, înregistrată pentru un eșanțon de volum $n=5$, s-au obținut: $\sum_i x_i^2 = 12.900$, $\sum_i x_i = 230$. Să se afle valoarea varianței.
- $\sigma^2 = 480$
 - $\sigma^2 = 520$
 - $\sigma^2 = 464$
6. Pentru o distribuție statistică după o variabilă X, înregistrată pentru un eșanțon de volum $n=8$, s-au obținut: $\sum_i x_i^2 n_i = 15.900$, $\sum_i x_i n_i = 350$. Să se afle valoarea abaterii medii pătratice.
- $\sigma = 9,52$
 - $\sigma = 8,57$
 - $\sigma = 6,53$
7. Pe un eșanțon de volum $n=30$ de angajați s-a înregistrat salariul obținut în luna octombrie 2004 și s-au obținut următoarele rezultate:
 $\sum_i x_i^2 = 900$, $\sum_i x_i = 150$ mil.lei. Intervalul mediu de variație este egal cu:
- (2,236; 2,764) mil. lei
 - (2,764; 7,236) mil. lei
 - (2,236; 5) mil. lei
8. Interpretarea corectă este
- Dacă $\beta_1 = 0$, atunci distribuția este asimetrică
 - Dacă $\beta_1 = 0$, atunci distribuția este simetrică
 - Dacă $\beta_1 > 0$, atunci distribuția este asimetrică

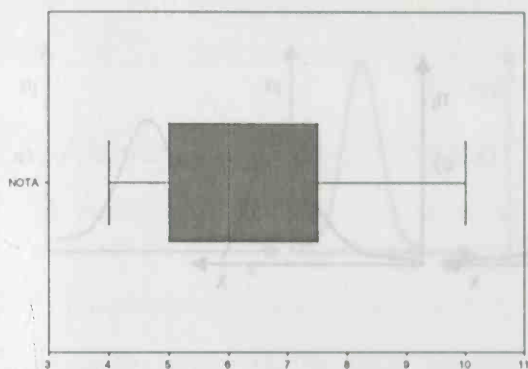
9. Dacă $Q_1 = 128,13$ milioane lei; $Me = 137,86$ milioane lei și $Q_3 = 147,92$ milioane lei, atunci coeficientul de asimetrie Yule este egal cu
- 141 și arată o distribuție simetrică
 - 0,067 și arată o distribuție moderat asimetrică
 - 0,067 și arată o distribuție puternic asimetrică
10. Dacă $M_0 = 137,5$ minute, $\bar{x} = 137,4$ minute și $\sigma = 12,74$ minute, atunci valoarea și interpretarea corectă a coeficientului empiric de asimetrie Pearson sunt
- 0,0078 și arată o asimetrie puternică la dreapta
 - 0,0078 și arată o ușoară asimetrie la stânga
 - 0,0078 și arată o ușoară asimetrie la stânga
11. Pentru un eșantion de volum $n = 100$, s-au obținut următoarele rezultate: $\sum_i (x_i - \bar{x})^2 \cdot n_i = 16260$, $\mu_3 = -456.33$. Care este valoarea coeficientului de asimetrie Pearson și interpretarea sa corectă ?
- $\beta_1 = 0,0484$, cu $\mu_3 < 0$, arată o distribuție simetrică
 - $\beta_1 = -0,0484$, cu $\mu_3 < 0$, arată o distribuție asimetrică la stânga
 - $\beta_1 = 0,0484$, cu $\mu_3 < 0$, arată o distribuție asimetrică la stânga
12. Folosind diagrama box-and-wiskers, o distribuție statistică poate fi reprezentată grafic astfel:



Precizați ce exprimă diagrama de mai sus:

- o distribuție asimetrică la dreapta
- o distribuție asimetrică la stânga
- o distribuție simetrică

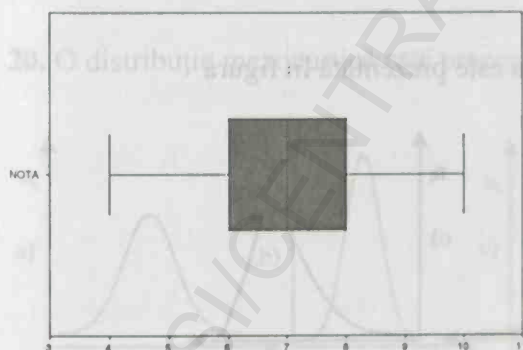
13. Folosind diagrama box-and-whiskers, o distribuție statistică poate fi reprezentată grafic astfel:



Precizați ce exprimă diagrama de mai sus:

- a) o distribuție asimetrică la dreapta
- b) o distribuție asimetrică la stânga
- c) o distribuție simetrică

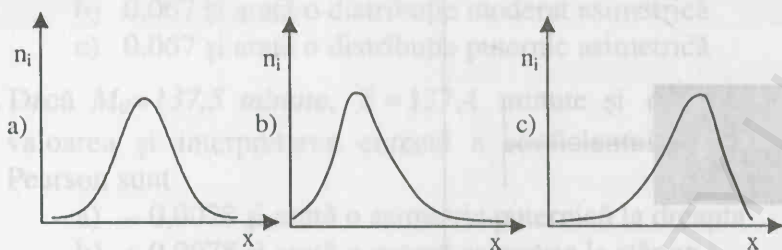
14. Folosind diagrama box-and-whiskers, o distribuție statistică poate fi reprezentată grafic astfel:



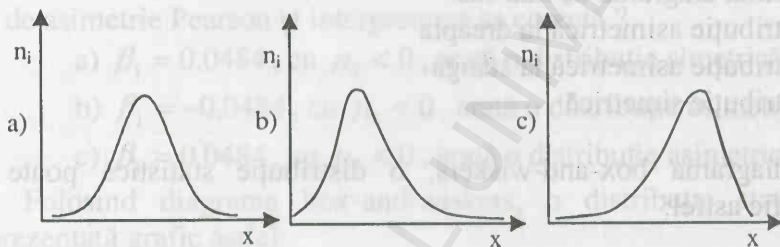
Precizați ce exprimă diagrama de mai sus:

- a) o distribuție asimetrică la dreapta
- b) o distribuție asimetrică la stânga
- c) o distribuție simetrică

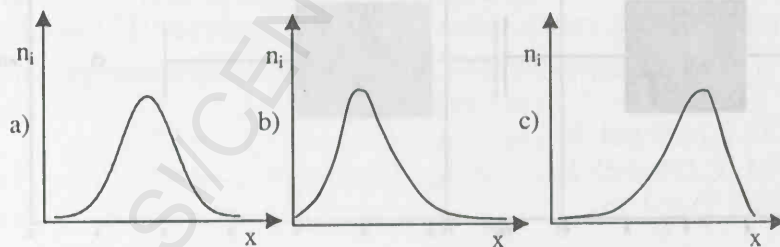
15. O distribuție simetrică este prezentată în figura



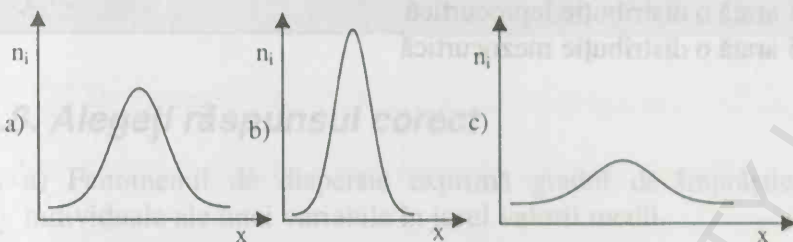
16. O distribuție asimetrică la dreapta este prezentată în figura



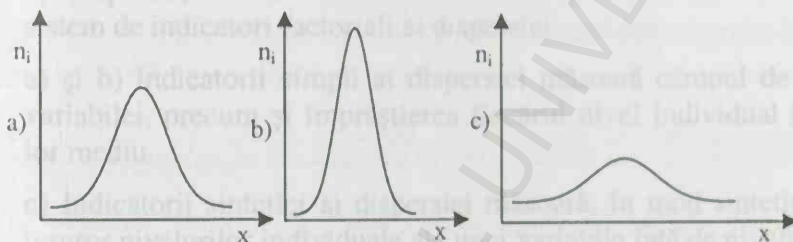
17. O distribuție asimetrică la stânga este prezentată în figura



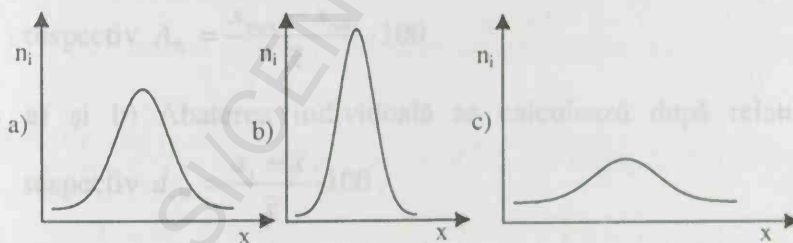
18. O distribuție leptocurtică este prezentată de figura



19. O distribuție platycurtică este prezentată de figura



20. O distribuție mezocurtică este prezentată de figura

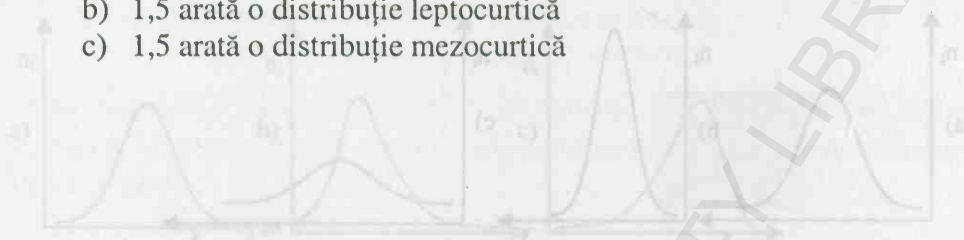


21. Dacă $\mu_4 = 54059,44$, $\mu_2 = 162,6$, atunci coeficientul de boltire Pearson și interpretarea sa corectă sunt

- a) $\beta_2 = 2,045$ arată o distribuție ușor platycurtică
- b) $\beta_2 = 2,045$ arată o distribuție leptocurtică
- c) $\beta_2 = 2,045$ arată o distribuție mezocurtică

22. Pentru un coeficient de boltire Pearson egal cu 4,5, coeficientul de exces Fisher și interpretarea sa corectă sunt

- 1,5 arată o distribuție platycurtică
- 1,5 arată o distribuție leptocurtică
- 1,5 arată o distribuție mezocurtică



16. O distribuție platycurtică este caracterizată de



17. O distribuție mezocurtică este caracterizată de



21. Dacă $k_1 = 24029,44$, $k_2 = 162,6$, atunci coeficientul de boltire Pearson

și interpretarea sa corectă sunt

- $k_1 = 2,042$ arată o distribuție ușor platycurtică
- $k_1 = 2,042$ arată o distribuție leptocurtică
- $k_1 = 2,042$ arată o distribuție mezocurtică



6.10. Răspunsuri corecte

6.8. Alegeți răspunsul corect

1. a) Fenomenul de dispersie exprimă gradul de împrăștiere a valorilor individuale ale unei variabile în jurul valorii medii.
2. a) Dispersia este datorată acțiunii factorilor aleatori.
3. a) și d) Dispersia, în cazul seriilor unidimensionale, este măsurată cu ajutorul unui sistem de indicatori simpli și sintetici ai dispersiei.
4. b) Dispersia, în cazul seriilor multivariate, este măsurată cu ajutorul unui sistem de indicatori factoriali ai dispersiei.
5. a) și b) Indicatorii simpli ai dispersiei măsoară câmpul de împrăștiere al variabilei, precum și împrăștierea fiecărui nivel individual față de nivelul lor mediu.
6. c) Indicatorii sintetici ai dispersiei măsoară, în mod sintetic, împrăștierea tuturor nivelurilor individuale ale unei variabile față de nivelul lor mediu.
7. a) și b) Indicatorii simpli ai dispersiei sunt amplitudinea variației și abaterea individuală.
8. a) și b) Amplitudinea variației se calculează după relația $A = x_{\max} - x_{\min}$, respectiv $A_{\%} = \frac{x_{\max} - x_{\min}}{\bar{x}} \cdot 100$.
9. a) și b) Abaterea individuală se calculează după relația $d_i = x_i - \bar{x}$, respectiv $d_{i\%} = \frac{x_i - \bar{x}}{\bar{x}} \cdot 100$.
10. a) și b) Indicatorii sintetici ai dispersiei sunt abaterea medie liniară și abaterea medie pătratică, varianța și coeficientul de variație.
11. b) și c) Abaterea medie liniară se calculează după relația $\bar{d} = \frac{\sum_i |x_i - \bar{x}| \cdot n_i}{n}$, respectiv $\bar{d} = \frac{\sum_i |x_i - \bar{x}|}{n}$.

12. a) și b) Varianța se calculează după relația $\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}$, respectiv

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2 \cdot n_i}{\sum n_i}.$$

13. a) și b) Abaterea medie pătratică se calculează după relația

$$\sigma = \sqrt{\frac{\sum d_i^2 \cdot n_i}{\sum n_i}}, \text{ respectiv } \sigma = \sqrt{\sigma^2}.$$

14. a) și b) Intervalul mediu de variație este definit de relația $\bar{x} \pm \bar{d}$, respectiv $\bar{x} \pm \sigma$.

15. a) Într-o distribuție normală standard, în intervalul mediu ($\bar{x} \pm \sigma$) apar 68,27% din cazuri.

16. b) și c) Coeficientul de variație se calculează după relația $v = \frac{\bar{d}}{\bar{x}} \cdot 100$, respectiv $v = \frac{\sigma}{\bar{x}} \cdot 100$.

17. a) Coeficientul de variație poate lua valori cuprinse între $0 < v < 100\%$.

18. a) și c) Varianța este egală cu $\sigma^2 = \overline{x^2} - (\bar{x})^2$, respectiv

$$\sigma^2 = \frac{1}{\sum n_i} \cdot \sum (x_i - a)^2 \cdot n_i - (\bar{x} - a)^2.$$

19. a) și b) Varianța este egală cu $\sigma^2 = \sigma_x^2 \cdot k^2$, unde $\sigma_x^2 = \frac{1}{n} \sum \left(\frac{x_i - \bar{x}}{k} \right)^2$,

$$\text{respectiv } \sigma_x^2 = \frac{1}{n} \sum \left(\frac{x_i - a}{k} \right)^2 \cdot k^2 - (\bar{x} - a)^2.$$

20. a) și b) Varianța unei caracteristici alternative este definită de relația

$$\sigma_p^2 = p \cdot q, \text{ respectiv, } \sigma_p^2 = p \cdot (1 - p).$$

21. a) și b) Indicatorii dispersiei în sistemul median sunt abaterea interquartilică, semiinterquartilia, coeficientul de variație interquartilic, precum și abaterea medie interdecilică și coeficientul de variație interdecilic.

22. b) Abaterea medie interquartilică se calculează după relația $A_Q = \frac{Q_3 - Q_1}{2}$.

23. c) Coeficientul de variație interquartilic se calculează după relația

$$V_Q = \left(\frac{Q_3 - Q_1}{Me} \right) \cdot 100.$$

24. a), b), c) și d) Indicatorii factoriali ai dispersiei sunt varianța de grupă (σ_j^2), media varianțelor de grupă ($\overline{\sigma^2}$), varianța mediilor de grupă față de media generală (δ^2) și varianța generală (σ^2).

25. a) Varianța de grupă măsoară variația sub influența factorilor aleatori (variația reziduală).

26. b) Varianța mediilor de grupă față de media generală (varianța intergrupe) măsoară variația sub influența factorilor esențiali.

27. c) Varianța generală măsoară variația sub influența factorilor întâmplători și esențiali.

28. a) și b) Sunt corecte relațiile $\sigma^2 = \overline{\sigma^2} + \delta^2$ și $\overline{\sigma^2} = \sigma^2 - \delta^2$.

29. a) și b) Varianța intragrupă (media varianțelor de grupă) se calculează după

relația $\overline{\sigma_p^2} = \frac{\sum_j p_j q_j n_{\cdot j}}{\sum_j n_{\cdot j}}$, $j = \overline{1, k}$, respectiv $\overline{\sigma^2} = \frac{\sum_j \sigma_j^2 n_{\cdot j}}{\sum_j n_{\cdot j}}$, unde

$$\sigma_j^2 = \frac{1}{n_{\cdot j}} \sum_i (x_i - \overline{x_j})^2 n_{ij}.$$

30. a) și b) Varianța intergrupe (varianța mediilor de grupă față de media

generală) se calculează după relația $\delta_p^2 = \frac{\sum_j (p_j - p)^2 n_{\cdot j}}{\sum_j n_{\cdot j}}$, respectiv

$$\delta^2 = \frac{\sum_j (\bar{x}_j - \bar{x})^2 n_{\cdot j}}{\sum_j n_{\cdot j}}.$$

31. a) și b) În cazul unei distribuții bidimensionale, cu ambele variabile exprimate cifric, se calculează media și varianța marginală, precum și medii și varianțe condiționate.

32. a) și b) Varianța condiționată (varianța de grupă) se determină pe baza

relației $\sigma_j^2 = \frac{1}{n_{\cdot j}} \sum_i (x_i - \bar{x}_j)^2 n_{ij}$, pentru variabila X, condiționată de $Y = y_j$,

respectiv $\sigma_i^2 = \frac{1}{n_{i\cdot}} \sum_j (y_j - \bar{y}_i)^2 n_{ij}$, pentru variabila Y, condiționată de $X = x_i$.

33. a), b) și c) Varianța generală se calculează după relațiile:

$$\sigma^2 = \frac{1}{n_{\cdot\cdot}} \sum_i (x_i - \bar{x})^2 n_{i\cdot}, \text{ cu } i = \overline{1, m}, \quad \sigma^2 = \frac{1}{n_{\cdot\cdot}} \sum_i x_i^2 n_{i\cdot} - (\bar{x})^2, \text{ cu } i = \overline{1, m},$$

$$\sigma^2 = \frac{1}{n_{\cdot\cdot}} \sum_j \sigma_j^2 n_{\cdot j} + \frac{1}{n_{\cdot\cdot}} \sum_j (\bar{x}_j - \bar{x})^2 n_{\cdot j}.$$

34. a) și b) Covarianța a două variabile, X și Y, este o măsură a covariației, respectiv a variației simultane a celor două variabile.

35. a) și b) Covarianța este definită de relația $\text{cov}(X, Y) = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{n}$,

$$i = \overline{1, n}, \text{ respectiv } \text{cov}(X, Y) = \frac{1}{n} \sum_i \sum_j (x_i - \bar{x})(y_i - \bar{y}) \cdot n_{ij}, \quad i = \overline{1, k}, \\ j = \overline{1, p}.$$

36. a) și b) Forma unei distribuții statistice unidimensionale poate fi apreciată cu ajutorul indicatorilor de asimetrie și boltire.

37. a) Indicatorii de asimetrie dau informații asupra modului de repartizare a frecvențelor de o parte sau alta a valorii centrale a unei serii.
38. b) Indicatorii de boltire exprimă măsura aglomerării frecvențelor în zona centrală, pe lângă medie.
39. a) Asimetria reprezintă deviația de la forma simetrică a distribuției.
40. b) și c) Aprecierea grafică a asimetriei poate fi făcută cu ajutorul curbei frecvențelor și a diagramei box-and-whiskers.
41. b) Într-o distribuție unidimensională simetrică, între cele trei mărimi medii fundamentale există relația $\bar{x} = Do = Me$, iar perechile de quantile Q_1 și Q_3 , respectiv D_1 și D_9 , sunt echidistante față de valoarea centrală.
42. a) și b) Asimetria se calculează după relația $As = \bar{x} - Mo$, respectiv $As = 3(\bar{x} - Me)$.
43. a) Coeficientul de asimetrie Yule (Bowley) se calculează după relația
- $$Cay = \frac{(q_2 - q_1)}{(q_2 + q_1)}.$$
44. a) Coeficientul empiric de asimetrie Pearson se calculează după relația
- $$Cas = \frac{\bar{x} - Do}{\sigma}$$
45. b) Coeficientul de asimetrie Pearson, calculat pe baza momentelor centrate, se determină după relația $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$.
46. a) O distribuție empirică, în comparație cu distribuția normală de aceeași medie și varianță, este leptocurtică atunci când prezintă o variație slabă a variabilei X însoțită de o variație puternică a frecvenței.
47. b) O distribuție empirică, în comparație cu distribuția normală de aceeași medie și varianță, este platycurtică atunci când prezintă o variație puternică a variabilei X însoțită de o variație slabă a frecvenței.
48. a) Coeficientul de boltire Pearson se calculează după relația $\beta_2 = \frac{\mu_4}{\mu_2^2}$.

49. a) și b) Coeficientul de boltire Fisher se calculează după relațiile:

$$\gamma_2 = \beta_2 - 3 ; \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3.$$

50. a) O distribuție este mezocurtică atunci când $\beta_2 = 3$, $\gamma_2 = 0$.

51. b) O distribuție este leptocurtică atunci când $\beta_2 > 3$, $\gamma_2 > 0$.

52. c) O distribuție este platicurtică atunci când $\beta_2 < 3$, $\gamma_2 < 0$.

6.9. Recunoașteți exemplul corect

1. b) Abaterea medie pătratică ($\sigma = 19$ milioane lei) arată că fiecare firmă din eșantionul studiat obține un profit lunar care se abate în medie, de la profitul mediu, cu 19 milioane lei, în sens pozitiv și negativ.

2. a) și c) Indicatorii dispersiei sunt: $A\% = 33,3\%$; $\bar{d} = 12$ mii lei; $\sigma = 14,142$ mii lei, $v = 11,78\%$. Arată un câmp de variație de 40 mii lei și o colectivitate omogenă, precum și un coeficient de variație de 11,78% și o medie reprezentativă.

Amplitudinea de variație se calculează după relația:

$$A = X_{\max} - X_{\min}$$

$$A = 140 - 100 = 40 \text{ mii lei}$$

Interpretare: Câmpul de variație al variabilei este de 40 mii lei. Valoarea coeficientului de variație ($v = 11,78\%$) tinde spre zero și arată că distribuția considerată se caracterizează printr-o dispersie mică, deci este o distribuție omogenă, iar media este reprezentativă.

3. c) Intervalul mediu de variație și coeficientul de variație au următoarele valori: (178,89 ; 184,61) milioane lei și 1,57%.

Intervalul mediu de variație este $(\bar{x} \pm \sigma)$,

$$\text{unde } \sigma = \sqrt{\sigma^2}$$

Abaterea medie pătratică este:

$$\sigma = \sqrt{8,1875} = 2,86$$

Prin urmare, intervalul mediu de variație este:

Coeficientul de variație se calculează după relația:

$$v = \frac{\sigma}{\bar{x}} \cdot 100$$

$$\bar{x} - \sigma = 181,75 - 2,86 = 178,89 \text{ mil lei}$$

$$\bar{x} + \sigma = 181,75 + 2,86 = 184,61 \text{ mil lei}$$

Coeficientul de variație se calculează astfel:

$$v = \frac{2,86}{181,75} \cdot 100 = 1,57\%$$

4. b) Rezultatele corecte sunt $\sigma = 12,74$ minute și $v = 9,27\%$.

Varianța se calculează după relația:

$$\sigma^2 = \frac{\sum_i (x_i - \bar{x})^2 n_i}{\sum_i n_i}$$

Abaterea medie pătratică este:

$$\sigma = \sqrt{\sigma^2}$$

Coeficientul de variație se calculează după relația:

$$v = \frac{\sigma}{\bar{x}} \cdot 100$$

Pentru exemplul considerat, varianța este:

$$\sigma^2 = \frac{16.224}{100} = 162,24$$

$$\sigma = \sqrt{162,24} = 12,74 \text{ minute}$$

Coeficientul de variație este:

$$v = \frac{12,74}{137,4} \cdot 100 = 9,27\%$$

5. c) Varianța este $\sigma^2 = 464$.

Varianța este egală cu diferența dintre media pătratelor și pătratul mediei:

$$\sigma^2 = \overline{x^2} - (\bar{x})^2 =$$

$$= \frac{\sum_i x_i^2}{n} - \left(\frac{\sum_i x_i}{n} \right)^2$$

Calcule:

$$\sigma^2 = \frac{12.900}{5} - \left(\frac{230}{5} \right)^2 = 2580 - 2116 = 464$$

6. b) Abaterea medie pătratică este $\sigma = 8,57$.

Varianța este egală cu diferența dintre media pătratelor și pătratul mediei:

$$\sigma^2 = \frac{15.900}{8} - \left(\frac{350}{8} \right)^2 = 1987,5 - 1914,06 = 73,44$$

$$\sigma^2 = \overline{x^2} - (\bar{x})^2 \quad \left| \quad \sigma = \sqrt{73,44} = 8,57 \right.$$

$$= \frac{\sum_i x_i^2 n_i}{\sum_i n_i} - \left(\frac{\sum_i x_i n_i}{\sum_i n_i} \right)^2$$

Abaterea medie pătratică este:

$$\sigma = \sqrt{\sigma^2}$$

7. b) Pe un eșantion de volum $n=30$ de angajați s-a înregistrat salariul obținut în luna octombrie 2004 și s-au obținut următoarele rezultate:

$$\sum_i x_i^2 = 900, \sum_i x_i = 150 \text{ mil. lei. Intervalul mediu de variație este:}$$

$(\bar{x} \pm \sigma)$, unde:

$$\bar{x} = \frac{\sum_i x_i}{n} = \frac{150}{30} = 5 \text{ mil. lei;}$$

$\sigma = \sqrt{\sigma^2}$. Varianța se calculează după relația:

$$\sigma^2 = \overline{x^2} - (\bar{x})^2 = \frac{\sum_i x_i^2}{n} - \left(\frac{\sum_i x_i}{n} \right)^2 = \frac{900}{30} - \left(\frac{150}{30} \right)^2 = 30 - 25 = 5. \quad \text{Abaterea}$$

medie pătratică este: $\sigma = \sqrt{\sigma^2} = \sqrt{5} = 2,236$.

Intervalul mediu de variație: $(\bar{x} \pm \sigma) = (5 \pm 2,236) = (2,764; 7,236)$ milioane lei.

8. b) și c) Dacă $\beta_1 = 0$, atunci distribuția este simetrică și dacă $\beta_1 > 0$, atunci distribuția este asimetrică.

9. b) Dacă $Q_1 = 128,13$; $Me = 137,86$ și $Q_3 = 147,92$, atunci coeficientul de asimetrie Yule este:

$$C_{as} = \frac{Q_1 - 2Me + Q_3}{Q_3 - Q_1} = \frac{147,92 - 2 \cdot 137,86 + 128,13}{147,92 - 128,13} = 0,067 \quad \text{și arată o}$$

distribuție moderat asimetrică.

10. b) Dacă $Mo = 137,5$ minute, $\bar{x} = 137,4$ minute și $\sigma = 12,74$ minute, atunci valoarea și interpretarea corectă a coeficientului empiric de

asimetrie Pearson sunt: $C_{as} = \frac{\bar{x} - Mo}{\sigma} = \frac{137,4 - 137,5}{12,74} = \frac{-0,1}{12,74} = -0,0078$

și arată o ușoară asimetrie la stânga.

11. c) Pentru un eșantion de volum $n = 100$, s-au obținut:

$$\sum_i (x_i - \bar{x})^2 \cdot n_i = 16260, \quad \mu_3 = -456,33. \text{ Valoarea coeficientului de}$$

asimetrie Pearson și interpretarea sa corectă sunt:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}, \text{ unde: } \mu_3 = \frac{\sum_i (x_i - \bar{x})^3 n_i}{\sum_i n_i}, \quad \mu_2 = \frac{\sum_i (x_i - \bar{x})^2 n_i}{\sum_i n_i}.$$

$$\mu_3^2 = (-456,33)^2 = 208237,1, \text{ iar } \mu_2^3 = \left(\frac{16260}{100}\right)^3 = (162,6)^3 = 4298942,3.$$

$$\beta_1 = \frac{208237,1}{4298942,3} = 0,0484 \text{ și arată o distribuție asimetrică la stânga } (\mu_3 < 0).$$

12. b) Folosind diagrama box-and-wiskers, o distribuție statistică poate fi reprezentată grafic astfel:

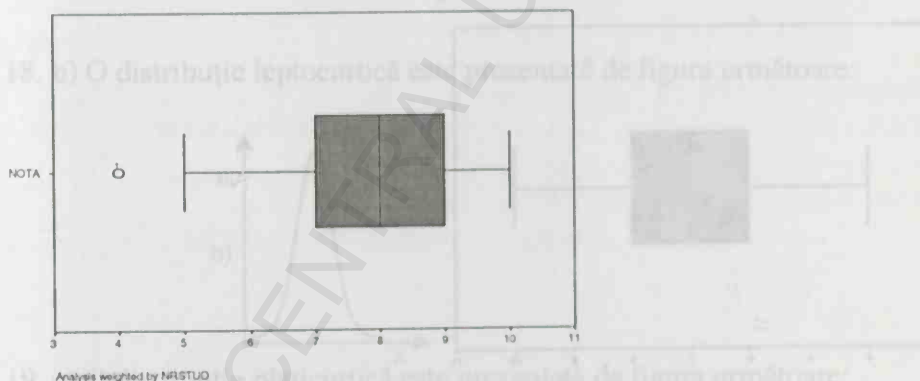


Diagrama de mai sus arată o distribuție asimetrică la stânga.

13. a) Folosind diagrama box-and-whiskers, o distribuție statistică poate fi reprezentată grafic astfel:

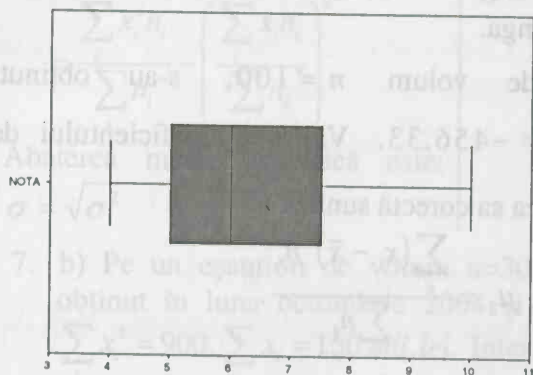


Diagrama de mai sus arată o distribuție asimetrică la dreapta.

14. c) Folosind diagrama box-and-whiskers, o distribuție statistică poate fi reprezentată grafic astfel:

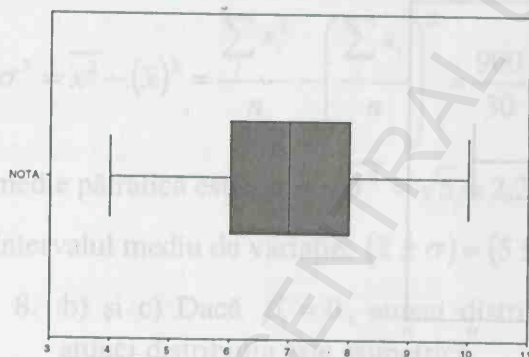
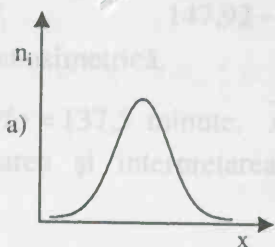
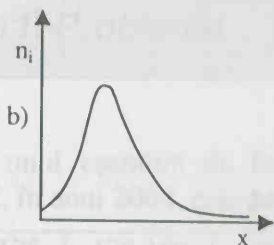


Diagrama de mai sus arată o distribuție simetrică.

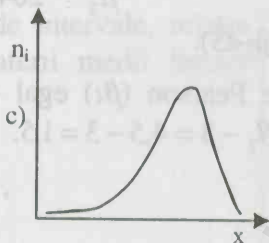
15. a) O distribuție simetrică este prezentată în figura de mai jos:



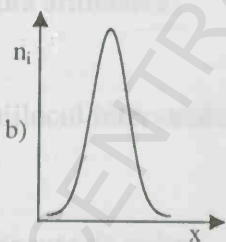
16. b) O distribuție asimetrică la dreapta este prezentată în figura de mai jos:



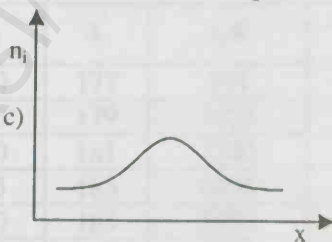
17. c) O distribuție asimetrică la stânga este prezentată în figura de mai jos:



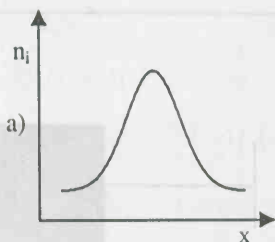
18. b) O distribuție leptocurtică este prezentată de figura următoare:



19. c) O distribuție platycurtică este prezentată de figura următoare:



20. a) O distribuție mezocurtică este prezentată de figura următoare:



21. a) Dacă $\mu_4 = 54059.44$, $\mu_2 = 162.6$, atunci coeficientul de boltire

Pearson și interpretarea sa corectă sunt: $\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{54059.44}{26438.76} = 2.045$ și

arată o distribuție ușor platycurtică ($\beta_2 < 3$).

22. b) Pentru un coeficient de boltire Pearson (β_1) egal cu 4,5, valoarea coeficientului Fisher este: $\gamma_2 = \beta_2 - 3 = 4.5 - 3 = 1.5$. Acesta arată o distribuție leptocurtică ($\gamma_2 > 0$).



6.11. Probleme

1. Distribuția unui eșantion de firme după cifra de afaceri lunară, X , în milioane lei, în anul 2004, este prezentată astfel:

x_{i-1}, x_i	176-178	178-180	180-182	182-184	184-186	186-188
n_i	20	25	40	35	30	10

Să se calculeze abaterea medie liniară.

Pentru o serie de intervale, relația de calcul a abaterii medii liniare este:

$$\bar{d} = \frac{\sum_i |x'_i - \bar{x}| n_i}{\sum_i n_i},$$

unde:

$$\bar{x} = \frac{\sum_i x'_i n_i}{\sum_i n_i}, \text{ media aritmetică}$$

$$x'_i = \frac{x_{i-1} + x_i}{2}, \text{ mijlocul intervalului}$$

Elementele de calcul sunt prezentate în tabelul 6.1.

$$\bar{x} = \frac{29080}{1600} = 181,75 \text{ mil. lei}$$

$$\bar{d} = \frac{387,50}{160} = 2,42 \text{ mil. lei}$$

Interpretare: Pe ansamblul firmelor se realizează o cifră de afaceri lunară care variază în medie față de nivelul mediu, cu 2,42 milioane lei, în sens pozitiv și negativ.

Tabelul 6.1. Elemente de calcul

x_{i-1}, x_i	n_i	x'_i	$x'_i n_i$	$x'_i - \bar{x}$	$ x'_i - \bar{x} $	$ x'_i - \bar{x} n_i$
176-178	20	177	354	-4,75	4,75	95,00
178-180	25	179	4475	-2,75	2,75	68,75
180-182	40	181	7240	-0,75	0,75	30,00
182-184	35	183	6405	1,25	1,25	43,75
184-186	30	185	5550	3,25	3,25	97,50
186-188	10	187	1870	5,25	5,25	52,50
Total	160	-	29080	-	-	387,50

2. Distribuția unui eșantion de firme după cifra de afaceri lunară, X , în milioane lei, în anul 2004, este:

x_{i-1}, x_i	176-178	178-180	180-182	182-184	184-186	186-188
n_i	20	25	40	35	30	10

Să se calculeze varianța (dispersia).

Pentru o serie de intervale, relația de calcul a varianței este:

$$\sigma^2 = \frac{\sum_i (x'_i - \bar{x})^2 n_i}{\sum_i n_i},$$

unde:

$$\bar{x} = \frac{\sum_i x'_i n_i}{\sum_i n_i}; \text{ media aritmetică}$$

$$x'_i = \frac{x_{i-1} + x_i}{2}, \text{ mijlocul intervalelor}$$

Elementele de calcul sunt prezentate în tabelul 6.2.

$$\bar{x} = \frac{29080}{160} = 181,75 \text{ mil. lei}$$

$$\sigma^2 = \frac{1310}{160} = 8,18$$

Tabelul 6.2. Elemente de calcul

x_{i-1}, x_i	n_i	x'_i	$x'_i n_i$	$x'_i - \bar{x}$	$ x'_i - \bar{x} $	$ x'_i - \bar{x} n$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^2$
176-178	20	177	354	-4,75	4,75	95,00	22,56	451,2
178-180	25	179	4475	-2,75	2,75	68,75	7,56	189,0
180-182	40	181	7240	-0,75	0,75	30,00	0,56	22,4
182-184	35	183	6405	1,25	1,25	43,75	1,56	54,6
184-186	30	185	5550	3,25	3,25	97,50	10,56	316,8
186-188	10	187	1870	5,25	5,25	52,50	27,56	275,6
Total	160	-	29080	-	-	387,50	-	1309,6

3. Distribuția unui eșantion de firme după cifra de afaceri lunară, X , în milioane lei, în anul 2004, este:

x_{i-1}, x_i	176-178	178-180	180-182	182-184	184-186	186-188
n_i	20	25	40	35	30	10

Să se calculeze valoarea abaterii medii pătratice.

Pentru o serie de intervale, relația de calcul a abaterii medii pătratice este:

$$\sigma = \sqrt{\sigma^2},$$

unde:

$$\sigma^2 = \frac{\sum_i (x_i' - \bar{x})^2 n_i}{\sum_i n_i}$$

$$\bar{x} = \frac{\sum_i x_i' n_i}{\sum_i n_i}$$

$$x_i' = \frac{x_{i-1} + x_i}{2}$$

Elementele de calcul sunt prezentate în tabelul 6.2.

$$\bar{x} = \frac{29080}{160} = 181,75 \text{ mil. lei}$$

$$\sigma^2 = \frac{1310}{160} = 8,18$$

$$\sigma = \sqrt{8,18} = 2,86 \text{ milioane lei}$$

Interpretare: Pe ansamblul firmelor se realizează o cifră de afaceri lunară care variază în medie față de nivelul mediu, cu 2,86 milioane lei, în sens pozitiv și negativ. Această valoare arată că aproximativ 68% din ansamblul firmelor obțin o cifră de afaceri cuprinsă în următorul interval de variație:
 $(\bar{x} \pm \sigma) = (181.75 \pm 2.86) = (178,89; 184,61)$ milioane lei.

4. Distribuția unui eșantion de firme după cifra de afaceri lunară, X, în milioane lei, în anul 2004, este:

x_{i-1}, x_i	176-178	178-180	180-182	182-184	184-186	186-188
n_i	20	25	40	35	30	10

Să se calculeze valoarea coeficientului de variație.

Pentru o serie de intervale, relația de calcul a coeficientului de variație este:

$$v = \frac{\sigma}{\bar{x}} \cdot 100,$$

Elementele de calcul sunt prezentate în tabelul 6.2.

$$\bar{x} = \frac{29080}{160} = 181,75 \text{ mil. lei}$$

unde:

$$\sigma = \sqrt{\sigma^2}$$

$$\sigma^2 = \frac{\sum_i (x_i' - \bar{x})^2 n_i}{\sum_i n_i}$$

$$\bar{x} = \frac{\sum_i x_i' n_i}{\sum_i n_i}$$

$$x_i' = \frac{x_{i-1} + x_i}{2}$$

$$\sigma^2 = \frac{1310}{160} = 8,18$$

$$\sigma = \sqrt{8,18} = 2,86 \text{ milioane lei}$$

$$v = \frac{2,86}{181,75} \cdot 100 = 1,57\%$$

Interpretare: Valoarea obținută pentru coeficientul de variație, ($v=1,57\%$)<17% arată o dispersie mică, respectiv, o colectivitate omogenă și o medie reprezentativă.

5. Distribuția muncitorilor dintr-o secție după timpul consumat pentru realizarea unui produs este caracterizată prin următoarele valori: $Q_1 = 128,13$ minute; $Q_2 = 137,86$ minute; $Q_3 = 147,92$ minute. Să se calculeze valoarea coeficientului de variație interquartilic.

Coeficientul de variație interquartilic se calculează după relația:

$$V_Q = \frac{A_Q}{Me} \cdot 100,$$

unde:

$$A_Q = \frac{Q_3 - Q_1}{2}$$

$$Me = Q_2$$

$$A_Q = \frac{147,92 - 128,13}{2} = 9,895$$

$$V_Q = \frac{9,895}{137,86} \cdot 100 = 7,17\%$$

Interpretare: Valoarea obținută pentru coeficientul de variație interquartilic, ($V_Q=7,17\%$)<17%, arată o dispersie mică, respectiv o colectivitate omogenă și o medie reprezentativă.

6. Distribuția muncitorilor dintr-o secție după timpul consumat pentru realizarea unui produs este caracterizată prin următoarele valori: $D_1=68$ minute; $D_5=53$ minute; $D_9=92$ minute. Să se calculeze coeficientul de variație interdecilic.

Coeficientul de variație interdecilic se calculează după relația:

$$V_D = \frac{A_D}{Me} \cdot 100,$$

unde:

$$A_D = \frac{D_9 - D_1}{2}$$

$$Me = D_5$$

$$A_D = \frac{92 - 68}{2} = 12$$

$$V_D = \frac{12}{53} \cdot 100 = 22,64\%$$

Interpretare:

Valoarea coeficientului de variație interdecilic, $17\% < (V_D = 22,64\%) < 35\%$, arată o dispersie mică, respectiv o colectivitate omogenă și o medie moderat reprezentativă.

7. Distribuția întreprinderilor din două ramuri, A și B, după termenul de dare în folosință a investițiilor este caracterizată prin următoarele valori: $n_A = 140$; $n_B = 270$, $\bar{x}_A = 4 \text{ ani}$, $\bar{x}_B = 6 \text{ ani}$, $\sigma_A^2 = 4,57$, $\sigma_B^2 = 5,33$. Să se măsoare influența factorilor întâmplători pe ansamblul întreprinderilor asupra variației termenului de dare în folosință.

Influența factorilor întâmplători pe total întreprinderi asupra variației termenului de dare în folosință se măsoară prin media varianțelor de grupă, calculată după relația:

$$\sigma^2 = \frac{\sum_j \sigma_j^2 n_j}{\sum_j n_j}$$

$$\bar{\sigma} = \sqrt{\sigma^2}$$

$$\sigma^2 = \frac{4,57 \cdot 140 + 5,33 \cdot 270}{410} = 5,07$$

$$\bar{\sigma} = \sqrt{5,07} = 2,25 \text{ ani}$$

Interpretare:

Influența factorilor întâmplători determină pe ansamblul întreprinderilor considerate o variație a termenului de dare în folosință egală cu $\pm 2,25$ ani

8. Distribuția întreprinderilor din două ramuri, A și B, după termenul de dare în folosință a investițiilor este caracterizată prin următoarele valori: $n_A = 140$; $n_B = 270$, $\bar{x}_A = 4 \text{ ani}$, $\bar{x}_B = 6 \text{ ani}$, $\sigma_A^2 = 4,57$, $\sigma_B^2 = 5,33$. Să se măsoare influența factorilor esențiali pe ansamblul întreprinderilor asupra variației termenului de dare în folosință.

Influența factorilor esențiali pe total întreprinderi asupra variației termenului de dare în folosință se măsoară prin varianța mediilor de grupă față de media generală, calculată după relația:

$$\delta^2 = \frac{\sum_j (\bar{x}_j - \bar{x})^2 \cdot n_{\cdot j}}{\sum_j n_{\cdot j}},$$

unde:

$$\bar{x}_j = \frac{\sum_i x_i n_{ij}}{n_{\cdot j}}, \text{ medii de grupă}$$

$$\bar{x} = \frac{\sum_j \bar{x}_j n_{\cdot j}}{\sum_j n_{\cdot j}}, \text{ media generală}$$

Media generală este:

$$\bar{x} = \frac{4 \cdot 140 + 6 \cdot 270}{410} = 5,32 \text{ ani}$$

Varianța mediilor de grupă față de media generală este:

$$\delta^2 = \frac{(4 - 5,32)^2 \cdot 140}{410} + \frac{(6 - 5,32)^2 \cdot 270}{410} = 0,8997$$

$$\delta = \sqrt{0,8997} = 0,95 \text{ ani}$$

Interpretare: Influența factorilor esențiali determină pe ansamblul întreprinderilor considerate o variație a termenului de dare în folosință egală cu $\pm 0,95$ ani.

9. Distribuția întreprinderilor din două ramuri, A și B, după termenul de dare în folosință a investițiilor este caracterizată prin următoarele valori: $n_A = 140$; $n_B = 270$, $\bar{x}_A = 4 \text{ ani}$, $\bar{x}_B = 6 \text{ ani}$, $\sigma_A^2 = 4,57$, $\sigma_B^2 = 5,33$. Să se măsoare influența factori esențiali și aleatori pe ansamblul întreprinderilor asupra variației termenului de dare în folosință.

Influența factorilor esențiali și aleatori pe ansamblul întreprinderilor asupra variației termenului de dare în folosință se măsoară prin varianța totală, calculată după relația:

$$\sigma^2 = \delta^2 + \overline{\sigma^2}, \quad \sigma = \sqrt{\sigma^2},$$

unde:

Media varianțelor de grupă este:

$$\overline{\sigma^2} = \frac{4,57 \cdot 140 + 5,33 \cdot 270}{140 + 270} = \frac{2078,9}{410} = 5,07$$

$$\bar{x} = \frac{4 \cdot 140 + 6 \cdot 270}{140 + 270} = \frac{560 + 1620}{410} = 5,32 \text{ a}$$

Varianța intergrupe este:

$$\delta^2 = \frac{(4 - 5,3)^2 \cdot 140 + (6 - 5,3)^2 \cdot 270}{410} = 0,89$$

$$\sigma^2 = \frac{\sum_j (\bar{x}_j - \bar{x})^2 n_{\cdot j}}{\sum_j n_{\cdot j}}$$
 este varianța

intergrupe. Arată influența factorilor esențiali.

$$\sigma^2 = \frac{\sum_j \sigma_j^2 \cdot n_{\cdot j}}{\sum_j n_{\cdot j}}$$
 este media

varianțelor de grupă. Arată influența factorilor aleatori.

$$\bar{x} = \frac{\sum_j \bar{x}_j \cdot n_{\cdot j}}{\sum_j n_{\cdot j}}$$
 este media pe total.

Varianța generală este:

$$\sigma^2 = 5,07 + 0,89 = 5,96$$

$$\sigma = \sqrt{5,96} = 2,44 \text{ ani}$$

Interpretare:

Influența factorilor esențiali și aleatori determină pe ansamblul întreprinderilor o variație a termenului de dare în folosință egală cu $\pm 2,44$ ani.

10. Distribuția muncitorilor firmei A după producția realizată, X, (piese) și salariul obținut, Y, (în USD), în luna ianuarie 2004, este:

$y_{j-1}, y_j \backslash x_{i-1}, x_i$	200-400	400-600	600-800	800-1000	$n_{i.}$
20-30	2	-	-	-	2
30-40	1	1	5	-	7
40-50	-	5	7	10	22
$n_{.j}$	3	6	12	10	31

Să se calculeze salariul mediu al muncitorilor care realizează o producție cuprinsă între 30-40 piese/lună.

Salariul mediu pentru grupa de muncă care realizează o producție cuprinsă între 30-40 piese/lună se calculează ca o medie condiționată, după relația:

$$\bar{y}_{j/i} = \frac{300 \cdot 1 + 500 \cdot 1 + 700 \cdot 5 + 900 \cdot 0}{7} = 614,28 \text{ USD}$$

$$\bar{y}_{j/i} = \frac{\sum_j y'_{ij} \cdot n_{ij}}{n_{i\cdot}}$$

$$y'_j = \frac{y_{j-1} + y_j}{2}$$

Interpretare:

Salariul mediu al muncitorilor care realizează o producție cuprinsă între 30-40 piese/lună este de 624,28 USD.

11. Distribuția muncitorilor firmei A după producția realizată, X, (piese) și salariul obținut, Y, (în USD), în luna ianuarie 2004, este:

y_{j-1}, y_j x_{i-1}, x_i	200-400	400-600	600-800	800-1000	$n_{i\cdot}$
20-30	2	-	-	-	2
30-40	1	1	5	-	7
40-50	-	5	7	10	22
$n_{\cdot j}$	3	6	12	10	31

Să se calculeze producția medie realizată de muncitorii care obțin un salariu lunar cuprins între 600-800 USD.

Producția medie realizată de muncitorii care obțin un salariu lunar cuprins între 600-800 USD se calculează ca o medie condiționată, după relația:

$$\bar{x}_{i/j} = \frac{\sum_i x'_i n_{ij}}{n_{\cdot j}}$$

$$x'_i = \frac{x_{i-1} + x_i}{2}$$

$$\bar{x}_{i/j} = \frac{25 \cdot 0 + 35 \cdot 5 + 45 \cdot 7}{12} = 40,83 \text{ piese}$$

Interpretare: Producția medie realizată de muncitorii care obțin un salariu lunar cuprins între 600-800 USD este de 40,83 piese ~ 41 piese.

12. Pentru o distribuție s-au obținut următoarele rezultate: $Q_1 = 150$; $M_c = 130$ și $Q_3 = 140$. Să se calculeze valoarea coeficientului de asimetrie Yule.

Coeficientul de asimetrie Yule se calculează după relația:

$$C_{as} = \frac{q_2 - q_1}{q_2 + q_1},$$

unde:

$$q_1 = Me - Q_1$$

$$q_2 = Q_3 - Me$$

$$q_1 = 130 - 150 = -20$$

$$q_2 = 140 - 130 = 10$$

$$C_{as} = \frac{10 + 20}{10 - 20} = -3$$

Interpretare:

Valoarea obținută, $C_{as} = -3 < 0$, arată că distribuția considerată este asimetrică la stânga.

13. Pentru o distribuție s-au obținut următoarele rezultate: $M_0 = 125$ minute, $\bar{x} = 130$ minute și $\sigma = 12$ minute. Să se calculeze valoarea coeficientului empiric de asimetrie Pearson.

Coeficientul empiric de asimetrie Pearson se determină după relația:

$$C_{as} = \frac{\bar{x} - M_0}{\sigma}$$

$$C_{as} = \frac{130 - 125}{12} = 0,417$$

Interpretare:

Valoarea obținută, $C_{as} = 0,417 > 0$, arată că distribuția considerată este puternic asimetrică la dreapta.

14. Pentru o distribuție s-au obținut următoarele rezultate: $\mu_4 = 40000$, $\mu_2 = 100$. Să se calculeze valoarea coeficientului de boltire Pearson.

Coeficientul de boltire Pearson se determină după relația:

$$\beta_2 = \frac{\mu_4}{\mu_2^2},$$

unde:

$$\mu_4 = 40000$$

$$\mu_2^2 = (100)^2 = 10000$$

$$\beta_2 = \frac{40000}{10000} = 4$$

$$\mu_4 = \frac{\sum_i (x_i - \bar{x})^4 n_i}{\sum_i n_i}$$

$$\mu_2 = \frac{\sum_i (x_i - \bar{x})^2 n_i}{\sum_i n_i}$$

Interpretare:

Pentru $(\beta_2 = 4) > 3$, distribuția considerată este leptocurtică.

15. Pentru o distribuție statistică s-a obținut o valoare a coeficientului de boltire Pearson egală cu 2,5. Să se calculeze coeficientul Fisher.

Coeficientul de boltire Fisher se determină după relația:

$$\gamma_2 = \beta_2 - 3$$

$\beta_2 = 2,5$ deci coeficientul de boltire Fisher este:

$$\gamma_2 = 2,5 - 3 = -0,5$$

Interpretare:

Pentru $(\gamma_2 = -0,5) < 0$, distribuția considerată este platicurtică.

Capitolul 7

INDICATORI AI CONCENTRĂRII

- ◆ Noțiuni
- ◆ Procedee grafice de determinare a concentrării (curba de concentrare)
- ◆ Procedee numerice de determinare a concentrării (abaterea medială-mediană, coeficientul de concentrare, coeficientul abaterii medii Gini, indicele de concentrare Gini)
- ◆ Alegeți răspunsul corect
- ◆ Recunoașteți exemplul corect
- ◆ Răspunsuri corecte
- ◆ Probleme

7.1. Noțiuni

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 185-186)

- Prin **concentrare** se exprimă *aglomerarea unităților unei colectivități sau a valorilor globale ale unei distribuții în jurul unei valori (de exemplu, a valorii centrale) a caracteristicii de grupare.*
- **Condiții pentru analiza distribuțiilor statistice cu ajutorul concentrării:**
 - însumarea variabilei de distribuție să aibă sens;
 - împărțirea valorii globale a variabilei între unitățile colectivității să fie posibilă.

7.2. Procedee grafice de determinare a concentrării (curba de concentrare)

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 187-189)

- **Curba de concentrare**
 - este reprezentarea grafică a variabilei „q” în funcție de variabila „p”. Valorile celor două variabile sunt definite de relațiile:

$$p_i = \frac{N_i}{\sum_i n_i}, \text{ cu } i = \overline{1, n}, \text{ unde } N_i = N_{i-1} + n_i;$$

$$q_i = \frac{L_i}{\sum_i x_i n_i}, \text{ cu } i = \overline{1, n}, \text{ unde } L_i = L_{i-1} + x_i n_i.$$

unde: p_i – efectivele relative cumulate până la nivelul „i” al caracteristicii de grupare,

q_i – valorile globale relative cumulate până la același nivel „i” al caracteristicii de grupare. Indicatorii variației absolute se obțin prin compararea, sub formă de scădere, a două nivele ale aceluiași indicator.

- **Grade de concentrare**

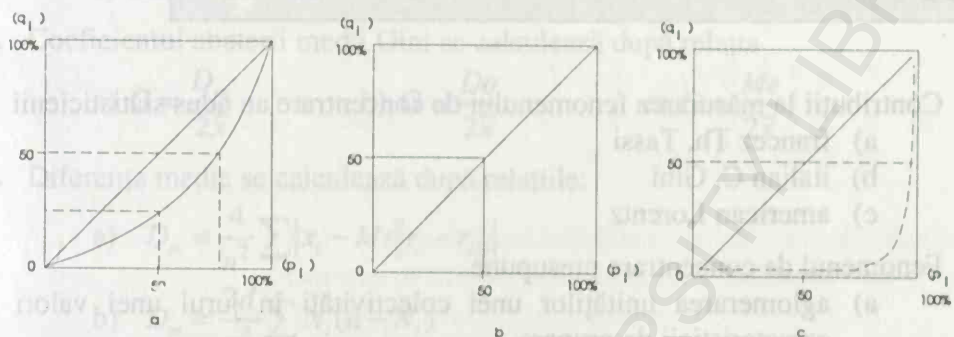


Fig. 7.1. Grade de concentrare: a) concentrare slabă; b) echirepartiție (lipsa concentrării); c) concentrare puternică

7.3. Procedee numerice de determinare a concentrării

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 190-197)

- Abaterea medială-mediană
- Coeficientul de concentrare
- Coeficientul abaterii medii Gini
- Indicele de concentrare Gini



7.4. Alegeți răspunsul corect

1. Contribuții la măsurarea fenomenului de concentrare au adus statisticienii
 - a) francez Th. Tassi
 - b) italian C. Gini
 - c) american Lorentz
2. Fenomenul de concentrare presupune
 - a) aglomerarea unităților unei colectivități în jurul unei valori a caracteristicii de grupare
 - b) aglomerarea caracteristicilor de grupare în jurul nivelului maxim al unităților statistice
 - c) aglomerarea valorilor globale în jurul unei valori a caracteristicii de grupare
3. Aprecierea concentrării implică studierea comparată a
 - a) efectivelor relative cumulate și a valorilor globale relative cumulate până la nivelul „i” al caracteristicii de grupare
 - b) structurii a două colectivități comparabile
 - c) structurii a două eșantioane extrase aleator
4. Măsurarea gradului de concentrare se poate realiza prin
 - a) abaterea medială-mediană
 - b) coeficienți ai gradului de concentrare
 - c) abaterea mediană-dominantă
 - d) curba Lorentz – Gini
5. Curba de concentrare se construiește în funcție de variabilele p și q . Valorile celor două variabile sunt definite de relațiile

$$\text{a) } p_i = \frac{N_i}{\sum_i n_i}, q_i = \frac{L_i}{\sum_i x_i n_i}, \text{ cu } i = \overline{1, n}$$

$$\text{b) } p_i = \frac{x_i}{\sum_i n_i}, q_i = \frac{X}{\sum_i x_i n_i}, \text{ cu } i = \overline{1, n}$$

$$\text{c) } p_i = \frac{x_i}{\sum_i n_i}, q_i = \frac{L_i}{\sum_{i=1}^k x_i n_i}, \text{ cu } i = \overline{1, n}$$

6. Abaterea medială – mediană se află după relația:

a) $\Delta Me = Do - Me$

b) $\Delta Mc = Do - Ml$

c) $\Delta Mc = Ml - Me$

7. Coeficientul abaterii medii Gini se calculează după relația

a) $G = \frac{D_m}{2\bar{x}}$

b) $G = \frac{Do}{2\bar{x}}$

c) $G = \frac{Me}{2\bar{x}}$

8. Diferența medie se calculează după relațiile:

a) $D_m = \frac{4}{n^2} \sum |x_i - Me| r_x - r_{Me}$

b) $D_m = \frac{2d}{n^2} \sum N_i(n - N_i)$

c) $D_m = \frac{2d}{\bar{x}} \sum N_i(n - N_i)$

9. Indicele de concentrare Gini se determină după relația

a) $i_G = 2 \cdot S_C$

b) $i_G = S_C$

c) $i_G = \frac{S_C}{2}$

10. Suprafața de concentrare se calculează prin

a) metoda trapezelor

b) metoda triunghiurilor

c) calcul integral

11. Suprafața de concentrare se calculează, prin metoda trapezelor, după relația

a) $S_C = \frac{1}{2} \left[1 - \sum_i (q_{i-1} + q_i) f_i \right]$

b) $S_C = \sum_i (p_i q_{i+1} - p_{i+1} q_i)$

c) $S_C = \frac{2d}{n^2} \sum_i N_i(n - N_i)$

12. Indicele de concentrare (i_G) se calculează, prin metoda triunghiurilor, după relația

$$a) i_G = \frac{1}{2} \left[1 - \sum_i (q_{i-1} + q_i) f_i \right]$$

$$b) i_G = \sum_i (p_i q_{i+1} - p_{i+1} q_i)$$

$$c) i_G = \frac{2d}{n^2} \sum_i N_i (n - N_i)$$

13. Curba de concentrare este folosită ca

- a) procedeu de aproximare a valorilor centrale
- b) procedeu de apreciere a gradului de concentrare
- c) procedeu de aproximare a asimetriei

14. Indicatorul exprimat prin relația $E_s = \sum_i s_i^2$, cu $i = \overline{1, N}$, reprezintă

- a) energia informațională Onicescu
- b) indicele Hirschman – Herfindahl
- c) diferența Hirschman

15. Coeficientul de concentrare Corrado Gini se calculează după relația

$$a) C_G = \sqrt{\sum_i g_i^2}, i = \overline{1, n} \text{ și ia valori cuprinse în intervalul } \left[\sqrt{\frac{1}{n}}, 1 \right]$$

$$b) C_G = \sqrt{\sum_i g_i^2}, i = \overline{1, n} \text{ și ia valori cuprinse în intervalul } [0, 1]$$

$$c) C_G = \sqrt{\sum_i g_i^2}, i = \overline{1, n} \text{ și ia valori cuprinse în intervalul } \left[1, \frac{1}{N} \right]$$

16. Indicatorii diversificării sunt

$$a) \text{ gradul de diversificare: } \frac{1}{N}$$

$$b) \text{ raportul de diversificare: } d = 1 - S_j, \text{ unde } S_j = \frac{n_j}{n}$$

$$c) \text{ indicele de diversificare: } i_D = 1 - \sum_{j=1}^k S_j^2, \text{ cu } S_j = \frac{n_j}{n}$$

17. În cazul variabilelor atributive, indicatorii de concentrare sunt

- a) energia informațională Onicescu: $E_S = \sum_{i=1}^N s_i^2$, $\sum_{i=1}^N s_i = 1$, cu valori

în intervalul $\left[\frac{1}{N}, 1\right]$

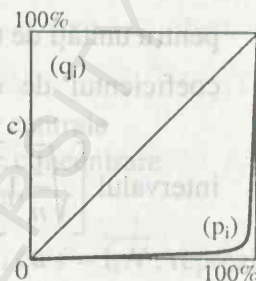
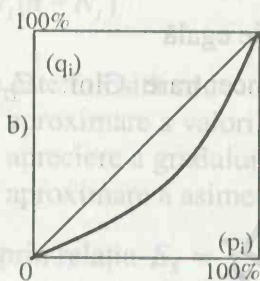
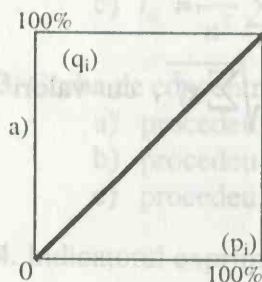
- b) diferența Hirschman: $H = H - H_e$, cu $H = \sum_{i=1}^N s_i^2$ și $H_e = H$ pentru unități de talie egală

- c) coeficientul de concentrare Gini: $C_G = \sqrt{\sum_{i=1}^N g_i^2}$, cu valori în intervalul $\left[\sqrt{\frac{1}{n}}, 1\right]$

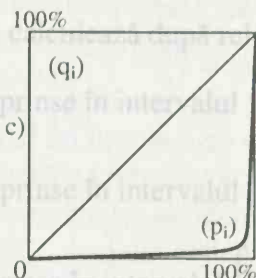
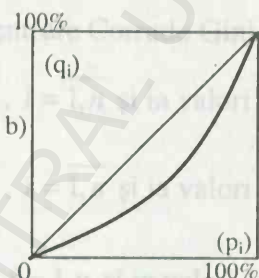
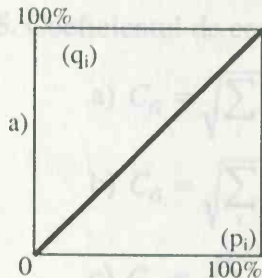


7.5. Recunoașteți exemplul corect

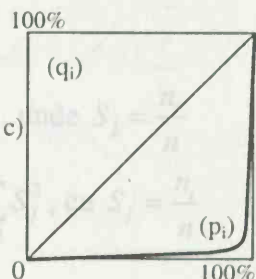
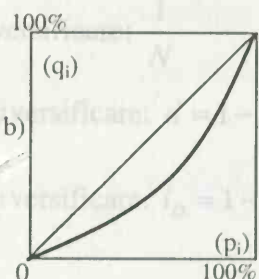
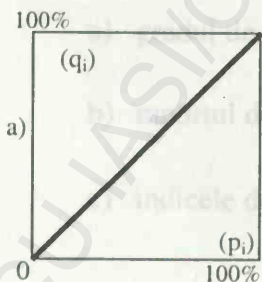
1. Care din diagramele de mai jos arată o echirepartiție?



2. Care din diagramele de mai jos arată o concentrare puternică?



3. Care din diagramele de mai jos arată o concentrare slabă?





7.6. Răspunsuri corecte

7.4. Alegeți răspunsul corect

1. b) și c) Contribuții la măsurarea fenomenului de concentrare au adus statisticienii C. Gini și Lorentz.
2. a) și c) Fenomenul de concentrare presupune aglomerarea unităților unei colectivități în jurul unei valori a caracteristicii de grupare sau aglomerarea valorilor globale în jurul unei valori a caracteristicii de grupare.
3. a) Aprecierea concentrării implică studierea comparată a efectivelor relative cumulate și a valorilor globale relative cumulate până la nivelul „ i ” al caracteristicii de grupare.
4. a), b) și d) Măsurarea gradului de concentrare se poate realiza prin abaterea medială – mediană, coeficienți ai gradului de concentrare și curba Lorentz – Gini.
5. a) Curba de concentrare se construiește în funcție de variabilele p și q .

Valorile celor două variabile sunt definite de relațiile $p_i = \frac{N_i}{\sum_i n_i}$,

$$q_i = \frac{L_i}{\sum_i x_i n_i}, \text{ cu } i = \overline{1, n}.$$

6. c) Abaterea medială – mediană se află după relația $\Delta Mc = Ml - Me$.

7. a) Coeficientul abaterii medii Gini se calculează după relația $G = \frac{D_m}{2\bar{x}}$.

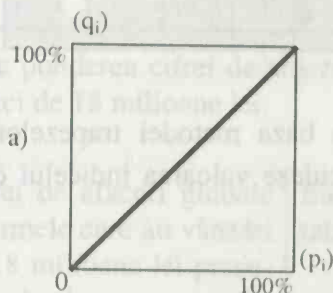
8. a) și b) Diferența medie se calculează după relațiile:

$$D_m = \frac{4}{n^2} \sum |x_i - Me| |r_x - r_{Me}| ; D_m = \frac{2d}{n^2} \sum N_i (n - N_i).$$

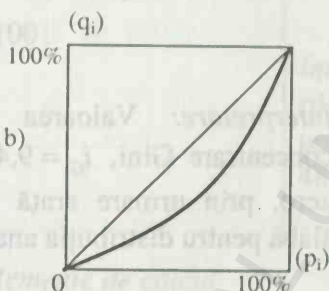
9. a) Indicele de concentrare Gini se determină după relația $i_G = 2 \cdot S_C$.
10. a), b) și c) Suprafața de concentrare se calculează prin metoda trapezelor, prin metoda triunghiurilor și prin calcul integral.
11. a) Suprafața de concentrare se calculează, prin metoda trapezelor, după relația $S_C = \frac{1}{2} \left[1 - \sum_i (q_{i-1} + q_i) f_i \right]$.
12. b) Indicele de concentrare (i_G) se calculează, prin metoda triunghiurilor, după relația $i_G = \sum_i (p_i q_{i+1} - p_{i+1} q_i)$.
13. a) și b) Curba de concentrare este folosită ca procedeu de aproximare a valorilor centrale și ca procedeu de apreciere a gradului de concentrare.
14. a) Indicatorul exprimat prin relația $E_S = \sum s_i^2$, cu $i = \overline{1, N}$, reprezintă energia informațională Onicescu.
15. a) Coeficientul de concentrare Corrado Gini se calculează după relația $C_G = \sqrt{\sum_i g_i^2}$, $i = \overline{1, n}$ și ia valori cuprinse în intervalul $\left[\sqrt{\frac{1}{n}}, 1 \right]$.
16. a), b) și c) Indicatorii diversificării sunt: gradul de diversificare, $\frac{1}{N}$, raportul de diversificare, $d = 1 - S_j$, unde $S_j = \frac{n_j}{n}$ și indicele de diversificare, $i_D = 1 - \sum_{j=1}^k S_j^2$, cu $S_j = \frac{n_j}{n}$.
17. a), b) și c) Toți indicatorii considerați sunt indicatori de concentrare pentru variabilele atributive.

7.5. Recunoașteți exemplul corect

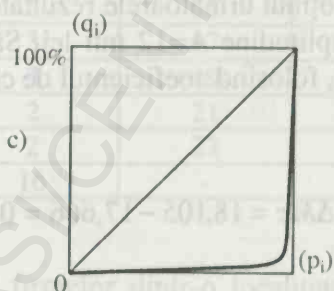
1. a) O echirepartiție este reprezentată în figura



2. b) O distribuție caracterizată printr-o concentrare slabă este reprezentată în figura



3. c) O distribuție caracterizată printr-o concentrare puternică este reprezentată în figura





7.7. Probleme

1. Pentru o distribuție statistică, pe baza metodei trapezelor, s-a obținut:

$$\sum_i (q_{i-1} + q_i) f_i = 0,906. \text{ Să se calculeze valoarea indicelui de concentrare}$$

Gini.

Indicele de concentrare Gini, pe baza metodei trapezelor, se calculează după relația:

$$i_G = 1 - \sum_i (q_{i-1} + q_i) f_i,$$

unde:

$$q_i = \frac{L_i}{\sum_i x_i n_i} \cdot 100, \quad L_i = \sum_{h=1}^i x_h n_h$$

$$f_i = \frac{n_i}{\sum_i n_i}$$

$$i_G = 1 - 0,906 = 0,094$$

$$\text{sau } i_G = 9,4\%$$

Interpretare: Valoarea indicelui de concentrare Gini, $i_G = 9,4\%$ tinde spre zero, prin urmare arată o concentrare slabă pentru distribuția analizată.

2. Pentru o distribuție statistică s-au obținut următoarele rezultate: $Me=17,666$ mii lei, $Ml=18,105$ mii lei și o amplitudine $A=12$ mii lei. Să se aprecieze gradul de concentrare al distribuției, folosind coeficientul de concentrare.

Coeficientul de concentrare se calculează după relația:

$$\Delta M_{\%} = \frac{\Delta Mc}{A_x} \cdot 100,$$

unde

$$\Delta Mc = Ml - Me$$

$$\Delta Mc = 18,105 - 17,666 = 0,439$$

$$\Delta M_{\%} = \frac{0,439}{12} \cdot 100 = 3,65\%$$

Interpretare: Valoarea coeficientului de concentrare ($\Delta M_{\%}=3,65\%$) tinde spre zero, prin urmare arată o concentrare slabă.

3. Distribuția firmelor dintr-o localitate după volumul vânzărilor (milioane lei), în luna ianuarie 2004, este prezentată în tabelul de mai jos.

x_{i-1}, x_i	12-14	14-16	16-18	18-20	20-22	22-24
n_i	1	5	3	3	2	2

Să se calculeze ponderea cifrei de afaceri globale realizate de firmele care au vânzări mai mici de 18 milioane lei.

Ponderea cifrei de afaceri globale realizate de firmele care au vânzări mai mici de 18 milioane lei presupune aflarea valorilor q_i corespunzătoare nivelului $X < (x_i = 18)$, după relația:

$$q_i = \frac{L_i}{\sum_i x_i n_i} \cdot 100,$$

unde:

$$L_i = \sum_h x_h n_h$$

Elementele de calcul sunt prezentate în tabelul 7.1.

$$q_i = \frac{139}{284} \cdot 100 = 48,94\%$$

Interpretare: Până la o valoare a vânzărilor de 18 milioane lei, corespunde o pondere a cifrei de afaceri globale de 48,94%.

Tabelul 7.1. Elemente de calcul

x_{i-1}, x_i	n_i	x_i	$x_i n_i$	L_i	$q_i (\%)$
12-14	1	13	13	13	4,57
14-16	5	15	75	88	30,98
16-18	3	17	51	139	48,94
18-20	3	19	57	196	69,01
20-22	2	21	42	238	83,80
22-24	2	23	46	284	100
Total	16	-	284	-	-

4. Distribuția firmelor dintr-o localitate după volumul vânzărilor (milioane lei), în luna ianuarie 2004, este prezentată în tabelul de mai jos.

x_{i-1}, x_i	12-14	14-16	16-18	18-20	20-22	22-24
n_i	1	5	3	3	2	2

Să se calculeze ponderea cifrei de afaceri globale realizate de firmele care au vânzări mai mici de 20 milioane lei.

Ponderea cifrei de afaceri globale realizate de firmele care au vânzări mai mici de 20 milioane lei presupune aflarea valorilor q_i corespunzătoare nivelului $X < (x_i = 20)$, după relația:

$$q_i = \frac{L_i}{\sum_i x_i n_i} \cdot 100$$

Elementele de calcul sunt prezentate în tabelul 7.1.

$$q_i = \frac{196}{284} \cdot 100 = 69,01\%$$

Interpretare: Până la o valoare a vânzărilor de 20 milioane lei, corespunde o pondere a cifrei de afaceri globale de 69,01%.

5. Pentru o distribuție statistică s-au obținut următoarele rezultate: $Dm=3,343$, $\sum_i x_i n_i = 284$, $\sum_i n_i = 16$. Să se aprecieze gradul de concentrare, folosind coeficientul abaterii medii Gini.

Coeficientul abaterii medii Gini se calculează după relația:

$$G = \frac{Dm}{2\bar{x}},$$

unde:

Dm este diferența medie

$$\bar{x} = \frac{\sum_i x_i n_i}{\sum_i n_i}, \text{ media aritmetică}$$

$$\bar{x} = \frac{284}{16} = 17,75$$

$$G = \frac{3,343}{2 \cdot 17,75} = 0,0941$$

Interpretare: Valoarea coeficientului abaterii medii Gini, $G=0,0941$, tinde spre zero, ceea ce arată o concentrare slabă.

6. Distribuția firmelor dintr-o localitate după volumul vânzărilor (milioane lei), în luna ianuarie 2004, este prezentată în tabelul de mai jos :

x_{i-1}, x_i	12-14	14-16	16-18	18-20	20-22	22-24
n_i	1	5	3	3	2	2

Să se calculeze abaterea medială-mediană.

Abaterea medială-mediană se calculează după relația:

$$\Delta Mc = Ml - Me,$$

unde:

$$Ml = x_{i-1} + d \cdot \frac{U^{Ml} - L_{i-1}}{x_i n_i}$$

$$U^{Ml} = \frac{\sum_i x_i n_i}{2}$$

$$Me = x_{i-1} + d \cdot \frac{U^{Me} - N_{i-1}}{n_i}$$

$$U^{Me} = \frac{\sum_i n_i}{2}$$

Când $n < 100$, atunci

$$U^{Me} = \frac{\left(\sum_i n_i \right) + 1}{2}$$

Elementele de calcul sunt prezentate în tabelul 7.2.

- calculul mediei

$$U^{Ml} = \frac{284}{2} = 142$$

$$(L_i = 196) > (U^{Ml} = 142)$$

Prin urmare, *intervalul medial* este intervalul (18-20).

$$Ml = 18 + 2 \cdot \frac{142 - 139}{57} = 18,105 \text{ mil. lei}$$

- calculul medianei

$$U^{Me} = \frac{16 + 1}{2} = 8,5$$

$$(N_i = 9) > (U^{Me} = 8,5)$$

Prin urmare, *intervalul median* este intervalul (16-18).

$$Me = 16 + 2 \cdot \frac{8,5 - 6}{3} = 17,666 \text{ mil. lei}$$

$$\Delta Mc = 18,105 - 17,666 = 0,439 \text{ mil. lei}$$

Interpretare:

Valoarea abaterii medială-mediană, $\Delta Mc = 0,439 \text{ mil. lei}$, tinde spre zero, ceea ce arată o concentrare slabă.

Tabelul 7.2. Elemente de calcul

x_{i-1}, x_i	n_i	N_i	x_i	$x_i n_i$	L_i
12-14	1	1	13	13	13
14-16	5	6	15	75	88
16-18	3	9	17	51	139
18-20	3	12	19	57	196
20-22	2	14	21	42	238
22-24	2	16	23	46	284
Total	16	-	-	284	-

7. Distribuția firmelor dintr-o localitate după volumul vânzărilor (milioane lei), în luna ianuarie 2004, este prezentată în tabelul de mai jos:

x_{i-1}, x_i	12-14	14-16	16-18	18-20	20-22	22-24
n_i	1	5	3	3	2	2

Să se calculeze coeficientul de concentrare medială-mediană ($\Delta M_{\%}$).

Coeficientul de concentrare medială-mediană se calculează după relația:

$$\Delta M_{(\%) } = \frac{\Delta Mc}{A_x} \cdot 100,$$

unde:

$$\Delta Mc = Ml - Me$$

$$A_x = x_{\max} - x_{\min}$$

$$Ml = x_{i-1} + d \cdot \frac{U^{Ml} - L_{i-1}}{x_i n_i}$$

$$U^{Ml} = \frac{\sum_i x_i n_i}{2}$$

$$Me = x_{i-1} + d \cdot \frac{U^{Me} - N_{i-1}}{n_i}$$

$$U^{Me} = \frac{\left(\sum_i n_i \right) + 1}{2}$$

Elementele de calcul sunt prezentate în tabelul 7.2.

Unitatea medială este:

$$U^{Ml} = \frac{284}{2} = 142$$

$$(L_i = 196) > (U^{Ml} = 142)$$

Intervalul medial este intervalul (18-20).

$$Ml = 18 + 2 \cdot \frac{142 - 139}{57} = 18,105 \text{ mil. lei}$$

Unitatea mediană este:

$$U^{Me} = \frac{16 + 1}{2} = 8,5$$

$$(N_i = 9) > (U^{Me} = 8,5)$$

Tabelul 7.4. Elemente de calcul

x_{i-1}, x_i	n_i	N_i	$N - N_i$
12-14	1	1	15
14-16	5	6	10
16-18	3	9	7
18-20	3	12	4
20-22	2	14	2
22-24	2	16	0
Total	16	16	0

Intervalul median este intervalul (16-18).

$$Me = 16 + 2 \cdot \frac{8,5 - 6}{3} = 17,666 \text{ mil. lei}$$

$$\Delta Mc = 18,105 - 17,666 = 0,439 \text{ mil. lei}$$

$$A_x = 24 - 12 = 12 \text{ mil. lei}$$

$$\Delta M_{(\%) } = \frac{0,439}{24 - 12} \cdot 100 = 3,65\%$$

Interpretare: Valoarea coeficientului abaterii medială-mediană,

$\Delta M_{(\%) } = 3,65\%$, tinde spre zero și arată o concentrare slabă.

8. Distribuția firmelor dintr-o localitate după volumul vânzărilor (milioane lei), în luna ianuarie 2004, este prezentată în tabelul de mai jos:

x_{i-1}, x_i	12-14	14-16	16-18	18-20	20-22	22-24
n_i	1	5	3	3	2	2

Să se calculeze valoarea coeficientului abaterii medii Gini.

Coeficientul abaterii medii Gini se calculează după relația:

$$G = \frac{Dm}{2\bar{x}},$$

unde:

$$Dm = \frac{2d}{n^2} \sum_i N_i(n - N_i)$$

$$\bar{x} = \frac{\sum_i x_i n_i}{\sum_i n_i}$$

Elementele de calcul sunt prezentate în tabelul 7.3.

$$\bar{x} = \frac{284}{16} = 17,75 \text{ mii lei}$$

$$Dm = \frac{2 \cdot 2}{16^2} \cdot 214 = 3,3437$$

$$G = \frac{3,3437}{2 \cdot 17,75} = 0,0942 \text{ sau } 9,42\%$$

Interpretare: Valoarea coeficientului abaterii medii Gini, $G=0,0942$, tinde spre zero și arată o concentrare slabă.

Tabelul 7.3. Elemente de calcul

x_{i-1}, x_i	n_i	N_i	$n - N_i$	$N_i(n - N_i)$
12-14	1	1	15	15
14-16	5	6	10	60
16-18	3	9	7	63
18-20	3	12	4	48
20-22	2	14	2	28
22-24	2	16	0	0
Total	16	-	-	214

9. Distribuția firmelor dintr-o localitate după volumul vânzărilor (milioane lei), în luna ianuarie 2004, este prezentată în tabelul de mai jos:

x_{i-1}, x_i	12-14	14-16	16-18	18-20	20-22	22-24
n_i	1	5	3	3	2	2

Să se calculeze valoarea indicelui de concentrare Gini, folosind metoda trapezelor.

Folosind metoda trapezelor, *indicele de concentrare Gini* se calculează după relația:

$$i_G = 1 - \sum_i (q_{i-1} + q_i) f_i,$$

unde:

$$q_i = \frac{L_i}{\sum_i x_i n_i}$$

$$f_i = \frac{n_i}{\sum_i n_i}$$

Elementele de calcul sunt prezentate în tabelul 7.4.

$$\sum_i (q_{i-1} + q_i) f_i = 0,905718$$

$$i_G = 1 - 0,905718 = 0,094282 \approx 0,0943$$

Interpretare: Valoarea indicelui de concentrare Gini, $i_G = 0,0943$, tinde spre zero și arată o concentrare slabă pentru distribuția considerată.

Tabelul 7.4. Elemente de calcul

x_i	n_i	f_i	q_i	$(q_{i-1}+q_i)$	$(q_{i-1}+q_i)f_i$
13	1	0,0625	0,0457	0,0457	0,002856
15	5	0,3125	0,3098	0,3555	0,111094
17	3	0,1875	0,4894	0,7992	0,149850
19	3	0,1875	0,6901	1,1795	0,221156
21	2	0,1250	0,8380	1,5281	0,191012
23	2	0,1250	1,0000	1,8380	0,229750
Total	16	1,00	-	-	0,905718

10. Distribuția firmelor dintr-o localitate după volumul vânzărilor (milioane lei), în luna ianuarie 2004, este prezentată în tabelul de mai jos:

x_{i-1}, x_i	12-14	14-16	16-18	18-20	20-22	22-24
n_i	1	5	3	3	2	2

Să se calculeze indicele de concentrare Gini, folosind metoda triunghiurilor.

Folosind metoda triunghiurilor, *indicele de concentrare Gini* se calculează după relația:

$$i_G = \sum_i (p_i q_{i+1} - p_{i+1} q_i),$$

unde:

$$q_i = \frac{L_i}{\sum_i x_i n_i}$$

$$p_i = \frac{N_i}{\sum_i n_i}$$

Elementele de calcul sunt prezentate în tabelul 7.5.

$$i_G = 0,09435$$

Interpretare: Valoarea indicelui de concentrare Gini, $i_G = 0,0943$, tinde spre zero și arată o concentrare slabă a distribuției analizate.

Tabelul 7.5. Elemente de calcul

x_i	n_i	p_i	q_i	$p_i q_{i+1}$	$p_{i+1} q_i$	$\frac{p_i q_{i+1} - p_{i+1} q_i}{2}$
13	1	0,0625	0,0457	0,0194	0,0171	0,00226
15	5	0,3750	0,3098	0,1835	0,1743	0,00924
17	3	0,5625	0,4894	0,3882	0,3670	0,02115
19	3	0,7500	0,6901	0,6285	0,6038	0,02470
21	2	0,8750	0,8380	0,8750	0,8380	0,03700
23	2	1,0000	1,0000	-	-	-
Total	16	-	-	-	-	0,09435

Capitolul 8

INDICATORI AI SERIILOR CRONOLOGICE

- ♦ Indicatori pentru caracterizarea nivelului și variației în timp
- ♦ Indicii dinamicii (indici agregați, indici calculați ca mărimi medii, indici calculați ca raport de medii)
- ♦ Calculul indicatorilor seriilor cronologice în SPSS
- ♦ Alegeți răspunsul corect
- ♦ Recunoașteți exemplul corect
- ♦ Răspunsuri corecte
- ♦ Probleme

8.1. Indicatori pentru caracterizarea nivelului și variației în timp

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economică, București, 2002, pp. 207-216)

- **Indicatori absoluți**
 - *Nivelul absolut*
 - *Volumul absolut*
 - *Sporul absolut* (modificarea absolută, creșterea sau descreșterea absolută): *sporul cu bază fixă* ($\Delta_{i/0}$) și *sporul cu bază mobilă* sau *în lanț* ($\Delta_{i/i-1}$).
- **Indicatori relativi**
 - *Ritmul sau indicele de variație*
 - *Ritmul sporului*
- **Indicatori medii**
 - *Nivelul mediu*
 - *Sporul mediu*
 - *Ritmul mediu al variației*
 - *Ritmul mediu al sporului*

8.2. Indicii dinamicii

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economică, București, 2002, pp. 216-224)

- **Categorii de indici**
 - *după natura variației*: indici ai dinamicii, indici teritoriali și indici calitativi;
 - *după baza de comparare*: indici cu bază fixă sau cu bază mobilă;
 - *după sfera de cuprindere a fenomenului*: indici individuali și indici de grup.

- **Indici agregați**
 - Indicii agregați se calculează prin raportarea nivelului agregat al unui fenomen din două perioade comparate: de bază (0) și curentă (1);
 - *Tipuri de indici agregați:* indicele Laspeyres, indicele Paasche, indicele Edgeworth, indicele Sidgwick-Drobisch.
- **Indici calculați ca mărimi medii**
 - *indicele mediu aritmetic*
 - *indicele mediu armonic*
 - *indicele mediu geometric – Irving Fischer*
- **Indici calculați ca raport de medii**
 - *indicele bifactorial*
 - *indicele cu structură fixă*
 - *indicele cu structură variabilă*

8.3. Calculul indicatorilor seriilor cronologice în SPSS

- **Demersul în SPSS**
 - Din meniul *Transform* se alege comanda *Compute*.
 - Se definește variabila considerată, de exemplu, sporul cu bază fixă sau mobilă, în *Target Variable*.
 - Se scrie relația de calcul în zona „*Numeric Expression*“. De exemplu, pentru calculul sporurilor cu bază fixă, avem relația:

$$\Delta_{1/0}^Y = y_{i1} - y_{i0}$$
 - Se folosesc variabilele Y_1 (nivelul fenomenului din perioada curentă), respectiv Y_0 (nivelul fenomenului din perioada de bază) din lista sursă, precum și operatorii aritmetici corespunzători.
 - Rezultatele se obțin în foaia „*Editorul de date*“.

• **Rezultatele în SPSS**

- Rezultatele în SPSS sunt prezentate astfel:

	regiune	an 1997	an 2000	spor	ritm
1	Nord-Est	14597,70	96348,40	81750,70	6,60
2	Sud-Est	14339,00	92868,20	78529,20	6,48
3	Sud	15511,40	98070,60	82559,20	6,32
4	Sud-Vest	10470,80	74888,90	64418,10	7,15
5	Vest	10572,80	75267,80	64695,00	7,12
6	Nord-Vest	12839,60	95009,80	82170,20	7,40
7	Centru	14429,50	101775,4	87345,90	7,05
8	București	15906,80	168791,7	152884,90	10,61



8.4. Alegeți răspunsul corect

1. O serie cronologică este denumită și
 - a) serie calitativă
 - b) serie dinamică
 - c) serie de spațiu
2. Sporul absolut este un
 - a) indicator absolut
 - b) indicator relativ
 - c) indicator mediu
 - d) indice de grup
3. Sporul absolut arată
 - a) de câte ori s-a modificat nivelul unui fenomen, într-o perioadă/moment, față de o altă perioadă/moment (bază de comparație)
 - b) cu cât s-a modificat nivelul unui fenomen, într-o perioadă/moment, față de o altă perioadă/moment (bază de comparație)
 - c) de câte ori s-a modificat în medie nivelul unui fenomen, într-o perioadă/moment, față de o altă perioadă/moment (bază de comparație)
4. Sporul cu bază fixă se determină după relația
 - a) $\Delta_{i/0}^y = y_i - y_0$
 - b) $\Delta_{i/i-1}^y = y_i - y_{i-1}$
 - c) $r_{n/0}^y = (y_n - y_0) / y_0$
5. Sporul cu bază mobilă se determină după relația
 - a) $\Delta_{i/0}^y = y_i - y_0$
 - b) $\Delta_{i/i-1}^y = y_i - y_{i-1}$
 - c) $r_{n/0}^y = (y_n - y_0) / y_0$
6. Indicatorii relativi se obțin prin
 - a) raportarea a două nivele comparabile
 - b) însumarea a două nivele
 - c) diferență între două nivele comparate

7. Mărimile absolute măsoară

- a) volumul unui fenomen și variația absolută
- b) variația relativă a nivelului unui fenomen
- c) variația medie a nivelului unui fenomen

8. În descompunerea pe factori a variației unui fenomen complex numărul de influențe se calculează după relația

- a) $N_K = 2^K - 1$
- b) $N_K = 2K - 1$
- c) $N_K = 2 - K$

9. Ritmul de variație arată

- a) de câte ori s-a modificat nivelul unui fenomen, într-o perioadă/ moment față de o altă perioadă/moment (bază de comparație)
- b) cu cât s-a modificat s-a modificat nivelul unui fenomen, într-o perioadă/moment față de o altă perioadă/moment (bază de comparație)
- c) cu cât s-a modificat în medie nivelul unui fenomen, într-o perioadă/moment față de o altă perioadă/moment (bază de comparație)

10. Creșterea relativă a nivelului unui fenomen într-o anumită perioadă / moment față de nivelul aceluiași fenomen din perioada / momentul de bază se măsoară prin

- a) sporul absolut
- b) ritmul creșterii
- c) ritmul variației

11. Valoarea absolută a unui procent de creștere se poate determina după relația

a) $\frac{\Delta_{i/0}}{r_{i/0}}$

b) $\frac{\Delta_{i/i-1}}{y_{i-1}}$

c) $\frac{\Delta_{i/0}}{y_0}$

12. Indicatorii medii ai seriilor cronologice măsoară

- a) nivelul fiecărui termen al seriei
- b) nivelul și variația individuală a termenilor seriei
- c) nivelul sau variația medie a tuturor termenilor seriei

13. Variația medie în timp a unui fenomen se poate calcula cu ajutorul

- a) sporului absolut
- b) ritmului mediu de variație
- c) sporului mediu al variației

14. Modificarea medie absolută a nivelului unui fenomen într-o perioadă poate fi măsurată prin
- ritmul mediu al variației
 - ritmul mediu al creșterii
 - sporul mediu al variației
15. Variația medie relativă a nivelului unui fenomen într-o perioadă poate fi măsurată prin
- ritmul mediu al variației
 - abaterea medie liniară
 - sporul mediu
16. Indicii statistici se calculează ca raport între două
- niveluri ale aceluiași fenomen, considerate la două momente diferite
 - valori corelate, la un moment dat
 - suma valorilor înregistrate pentru două variabile
17. Verificarea gradului și sensului în care sistemul de ponderare folosit influențează valoarea indicilor se efectuează pe baza relației
- Bortkiewicz
 - Edgeworth
 - Paasche
18. Pentru caracterizarea seriilor cronologice, sunt folosiți următorii indicatori exprimați în mărimi absolute
- nivelul absolut
 - volumul absolut
 - sporul absolut
 - ritmul de variație

19. Media cronologică se calculează după relațiile

$$a) \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

$$b) \bar{y}_{cr} = \frac{\frac{y_1}{2} + y_2 + \dots + y_{n-1} + \frac{y_n}{2}}{n-1}$$

$$c) \bar{y}_{cr} = \frac{\frac{y_1 + y_2}{2} t_1 + \frac{y_2 + y_3}{2} t_2 + \dots + \frac{y_{n-1} + y_n}{2} t_{n-1}}{t_1 + t_2 + \dots + t_{n-1}}$$

$$d) \bar{y} = \frac{\sum_i y_i t_i}{\sum_i t_i}$$

20. Ritmul mediu al variației se poate calcula după metoda
- mediei geometrice
 - mediei parabolice
 - trendului exponențial
 - mediilor mobile
21. Metoda indicilor statistici se utilizează pentru
- caracterizarea variației medii relative a fenomenelor în timp
 - măsurarea intensității fenomenelor în spațiu
 - măsurarea influenței factorilor asupra dinamicii fenomenelor complexe
 - măsurarea corelației dintre fenomene
22. Calculul indicilor agregați se efectuează după indicii:
- Laspeyres
 - Edgeworth
 - Sidgwick-Drobish
 - Pearson
 - Bortkiewicz
23. Indicele prețului de consum este folosit în calitate de
- mijloc de estimare a variației medii a prețurilor mărfurilor cumpărate și a tarifelor serviciilor utilizate de populație
 - mijloc de măsurare a inflației din sfera consumului și de determinare a puterii de cumpărare a veniturilor, a salariilor și a pensiilor
 - mijloc de fundamentare a calculului dobânzii reale, la recalcularea unor indicatori valorici din domeniul comerțului cu amănuntul
 - la negocierea salariilor, a pensiilor, a burselor, la fundamentarea deciziilor politice din domeniul social
 - la realizarea unor comparații internaționale
24. Descompunerea dinamicii unui fenomen complex se poate efectua prin
- procedeul substituiri în lanț
 - procedeul restului nedescompus
 - procedeul Lagrange
 - procedeul Struck

25. Indicii calculați ca raport a două medii măsoară

- variația bifactorială
- sporul mediu
- schimbarea structurală
- schimbarea nivelului caracteristicii



8.5. Recunoașteți exemplul corect

1. Indicele tip Laspeyres al prețului este

$$a) I_{1/0}^p = \frac{\sum_i p_{11} q_{i0}}{\sum_i p_{i0} q_{i0}}$$

$$b) I_{1/0}^p = \frac{\sum_i p_{11} q_{11}}{\sum_i p_{i0} q_{i1}}$$

$$c) I_{1/0}^p = \frac{\sum_i p_{11} q_{11}}{\sum_i p_{i0} q_{i0}}$$

2. Indicele tip Paasche al prețului este

$$a) I_{1/0}^p = \frac{\sum_i p_{11} q_{i0}}{\sum_i p_{i0} q_{i0}}$$

$$b) I_{1/0}^p = \frac{\sum_i p_{11} q_{11}}{\sum_i p_{i0} q_{11}}$$

$$c) I_{1/0}^p = \frac{\sum_i p_{11} q_{11}}{\sum_i p_{i0} q_{i0}}$$

3. Indicele tip Fischer al prețului este

$$a) I_{1/0}^p = \sqrt{\frac{\sum_i p_{11} q_{11}}{\sum_i p_{11} q_{i0}} \cdot \frac{\sum_i p_{i0} q_{11}}{\sum_i p_{i0} q_{i0}}}$$

$$b) I_{1/0}^p = \sqrt{\frac{\sum_i p_{11} q_{11}}{\sum_i p_{i0} q_{i0}} \cdot \frac{\sum_i p_{11} q_{11}}{\sum_i p_{i0} q_{11}}}$$

$$c) I_{1/0}^p = \sqrt{\frac{\sum_i p_{11} q_{i0}}{\sum_i p_{i0} q_{i0}} \cdot \frac{\sum_i p_{11} q_{11}}{\sum_i p_{i0} q_{11}}}$$

4. Indicele tip Sidgwick-Drobisch al prețului este

$$a) I_{1/0}^p = \frac{\sum_i p_{i1}q_{i0} + \sum_i p_{i1}q_{i1}}{\sum_i p_{i0}q_{i0} + \sum_i p_{i0}q_{i1}}$$

$$b) I_{1/0}^p = \frac{\frac{\sum_i p_{i1}q_{i0}}{\sum_i p_{i0}q_{i0}} + \frac{\sum_i p_{i1}q_{i1}}{\sum_i p_{i0}q_{i1}}}{2}$$

$$c) I_{1/0}^p = \sqrt{\frac{\sum_i p_{i1}q_{i0}}{\sum_i p_{i0}q_{i0}} \cdot \frac{\sum_i p_{i1}q_{i1}}{\sum_i p_{i0}q_{i1}}}$$

5. Indicele tip Laspeyres al volumului fizic al producției este

$$a) I_{1/0}^q = \frac{\sum_i p_{i1}q_{i1}}{\sum_i p_{i1}q_{i0}} \quad b) I_{1/0}^q = \frac{\sum_i p_{i0}q_{i1}}{\sum_i p_{i0}q_{i0}}$$

$$c) I_{1/0}^q = \frac{\sum_i p_{i1}q_{i1}}{\sum_i p_{i0}q_{i0}}$$

6. Indicele tip Paasche al volumului fizic al producției este

$$a) I_{1/0}^q = \frac{\sum_i p_{i1}q_{i1}}{\sum_i p_{i0}q_{i0}} \quad b) I_{1/0}^q = \frac{\sum_i p_{i0}q_{i1}}{\sum_i p_{i0}q_{i0}}$$

$$c) I_{1/0}^q = \frac{\sum_i p_{i1}q_{i1}}{\sum_i p_{i1}q_{i0}}$$

7. Indicele tip Fischer al volumului fizic al producției este

$$a) I_{1/0}^q = \sqrt{\frac{\sum_i p_{i1}q_{i1}}{\sum_i p_{i1}q_{i0}} \cdot \frac{\sum_i p_{i0}q_{i1}}{\sum_i p_{i0}q_{i0}}}$$

$$b) I_{1/0}^q = \sqrt{\frac{\sum_i p_{11} q_{11}}{\sum_i p_{10} q_{10}} \cdot \frac{\sum_i p_{11} q_{10}}{\sum_i p_{10} q_{11}}}$$

$$c) I_{1/0}^q = \sqrt{\frac{\sum_i p_{11} q_{11}}{\sum_i p_{10} q_{11}} \cdot \frac{\sum_i p_{11} q_{10}}{\sum_i p_{10} q_{10}}}$$

8. Indicele tip Sidgwick-Drobisch al volumului fizic este

$$a) I_{1/0}^q = \frac{\sum_i p_{11} q_{10} + \sum_i p_{11} q_{11}}{\sum_i p_{10} q_{10} + \sum_i p_{10} q_{11}}$$

$$b) I_{1/0}^q = \frac{\frac{\sum_i p_{11} q_{11}}{\sum_i p_{11} q_{10}} + \frac{\sum_i p_{10} q_{11}}{\sum_i p_{10} q_{10}}}{2}$$

$$c) I_{1/0}^q = \sqrt{\frac{\sum_i p_{11} q_{10}}{\sum_i p_{10} q_{10}} \cdot \frac{\sum_i p_{11} q_{11}}{\sum_i p_{10} q_{11}}}$$



8.6. Răspunsuri corecte

8.4. Alegeți răspunsul corect

1. b) O serie cronologică este denumită și serie dinamică.
2. a) Sporul absolut este un indicator absolut.
3. b) Sporul absolut arată cu cât s-a modificat nivelul unui fenomen, într-o perioadă/moment, față de o altă perioadă/moment (bază de comparație).
4. a) Sporul cu bază fixă se determină după relația $\Delta_{i/0}^y = y_i - y_0$.
5. b) Sporul cu bază mobilă se determină după relația $\Delta_{i/t-1}^y = y_i - y_{t-1}$.
6. a) Indicatorii relativi se obțin prin raportarea a două nivele comparabile.
7. a) Mărimile absolute măsoară volumul unui fenomen și variația absolută.
8. a) În descompunerea pe factori a variației unui fenomen complex, numărul de influențe se calculează după relația $N_K = 2^K - 1$.
9. a) Ritmul de variație arată de câte ori s-a modificat nivelul unui fenomen într-o anumită perioadă/moment, față de o altă perioadă/moment (bază de comparație).
10. b) Creșterea relativă a nivelului unui fenomen într-o anumită perioadă / moment, față de nivelul aceluiași fenomen din perioada/momentul de bază se măsoară prin ritmul creșterii.
11. a) Valoarea absolută a unui procent de creștere se poate determina după relația $\frac{\Delta_{i/0}}{r_{i/0}}$.
12. b) Indicatorii medii ai seriilor cronologice măsoară nivelul și variația medie a tuturor termenilor seriei.
13. b) și c) Variația medie în timp a unui fenomen se poate calcula cu ajutorul ritmului mediu și sporului mediu al variației.
14. c) Modificarea medie absolută a nivelului unui fenomen într-o perioadă poate fi măsurată prin sporul mediu al variației.

15. a) Variația medie relativă a nivelului unui fenomen într-o perioadă poate fi măsurată prin ritmul mediu al variației.
16. a) Indicii statistici se calculează ca raport între două niveluri ale aceluiași fenomen, la două momente diferite.
17. a) Verificarea gradului și sensului în care sistemul de ponderare folosit influențează valoarea indicilor se efectuează pe baza relației Bortkiewicz.
18. a), b) și c) Pentru caracterizarea seriilor cronologice, sunt folosiți următorii indicatori exprimați în mărimi absolute: nivelul absolut, volumul absolut și sporul absolut.
19. b) și c) Media cronologică se calculează după relațiile
- $$\bar{y}_{cr} = \frac{\frac{y_1}{2} + y_2 + \dots + y_{n-1} + \frac{y_n}{2}}{n-1};$$
- $$\bar{y}_{cr} = \frac{\frac{y_1 + y_2}{2} t_1 + \frac{y_2 + y_3}{2} t_2 + \dots + \frac{y_{n-1} + y_n}{2} t_{n-1}}{t_1 + t_2 + \dots + t_{n-1}}.$$
20. a), b) și c) Ritmul mediu al variației se poate calcula după metoda mediei geometrice, metoda mediei parabolice și metoda trendului exponențial.
21. a) și c) Metoda indicilor statistici se utilizează pentru caracterizarea variației medii relative a fenomenelor în timp și pentru măsurarea influenței factorilor asupra dinamicii fenomenelor complexe.
22. a), b) și c) Calculul indicilor agregati se efectuează după indicii Laspeyres, indicii Edgeworth și indicii Sidgwick-Drobish.
23. a), b), c), d) și e) Indicele prețului de consum este folosit în calitate de mijloc de estimare a variației medii a prețurilor mărfurilor cumpărate și a tarifelor serviciilor utilizate de populație, mijloc de măsurare a inflației din sfera consumului și de determinare a puterii de cumpărare a veniturilor, a salariilor și a pensiilor, mijloc de fundamentare a calculului dobânzii reale, la recalcularea unor indicatori valorici din domeniul comerțului cu amănuntul, la negocierea salariilor, a pensiilor, a burselor, la fundamentarea deciziilor politice din domeniul social și la realizarea unor comparații internaționale.

24. a), c) și d) Descompunerea dinamicii unui fenomen complex se poate efectua prin procedeul substituirii în lanț, procedeul restului nedescompus, procedeul Lagrange și procedeul Struck.

25. a), c) și d) Indicii calculați ca raport a două medii măsoară variația bifactorială, schimbarea structurală și schimbarea nivelului caracteristicii.

8.5. Recunoașteți exemplul corect

1. a) Indicele tip Laspeyres al prețurilor este $I_{1/0}^p = \frac{\sum_i p_{i1} q_{i0}}{\sum_i p_{i0} q_{i0}}$.

2. b) Indicele tip Paasche al prețului este $I_{1/0}^p = \frac{\sum_i p_{i1} q_{i1}}{\sum_i p_{i0} q_{i1}}$.

3. c) Indicele tip Fischer al prețului este $I_{1/0}^p = \sqrt{\frac{\sum_i p_{i1} q_{i0}}{\sum_i p_{i0} q_{i0}} \cdot \frac{\sum_i p_{i1} q_{i1}}{\sum_i p_{i0} q_{i1}}}$.

4. b) Indicele tip Sidgwick-Drobisch al prețului este

$$I_{1/0}^p = \frac{\frac{\sum_i p_{i1} q_{i0}}{\sum_i p_{i0} q_{i0}} + \frac{\sum_i p_{i1} q_{i1}}{\sum_i p_{i0} q_{i1}}}{2}.$$

5. b) Indicele tip Laspeyres al volumului fizic al producției este

$$I_{1/0}^q = \frac{\sum_i p_{i0} q_{i1}}{\sum_i p_{i0} q_{i0}}.$$

6. c) Indicele tip Paasche al volumului fizic al producției este

$$I_{1/0}^q = \frac{\sum_i p_{i1} q_{i1}}{\sum_i p_{i1} q_{i0}}.$$

7. a) Indicele tip Fischer al volumului fizic al producției este

$$I_{1/0}^q = \sqrt{\frac{\sum_i p_{11} q_{11}}{\sum_i p_{11} q_{10}} \cdot \frac{\sum_i p_{10} q_{11}}{\sum_i p_{10} q_{10}}}$$

8. b) Indicele tip Sidgwick-Drobisch al volumului fizic este

$$I_{1/0}^q = \frac{\frac{\sum_i p_{11} q_{11}}{\sum_i p_{11} q_{10}} + \frac{\sum_i p_{10} q_{11}}{\sum_i p_{10} q_{10}}}{2}$$



8.7. Probleme

1. Volumul exporturilor României înregistrat în perioada 1997-2001 este prezentat în tabelul de mai jos

Anii	1997	1998	1999	2000	2001
Valoare exporturi (milioane USD)	8431	8302	8487	10367	11385

Sursa : Anuarul statistic al României, C.N.S. 2002, p. 449

Să se calculeze cu cât s-a redus, în valoare absolută, volumul exporturilor României în anul 1998 față de anul 1997.

Valoarea absolută a reducerii volumului exporturilor României, în anul 1998 față de anul 1997, se află prin calculul sporului absolut cu bază fixă, după relația:

$$\Delta_{1/0} = y_1 - y_0$$

Sporul absolut cu bază fixă este:

$$\Delta_{1/0} = y_1 - y_0 = 8302 - 8431 = -129$$

Interpretare:

Volumul exporturilor României s-a redus, în anul 1998 față de anul 1997, cu 129 milioane USD.

2. Volumul exporturilor României înregistrat în perioada 1997-2001 este prezentat în tabelul de mai jos :

Anii	1997	1998	1999	2000	2001
Valoare exporturi (milioane USD)	8431	8302	8487	10367	11385

Sursa : Anuarul statistic al României, C.N.S. 2002, p. 449

Să se calculeze cu cât a crescut, în valoare absolută, volumul exporturilor României în anul 2001 față de anul 1997.

Valoarea absolută a creșterii volumului exporturilor României, în anul 2001 față de anul 2000, se află prin calculul sporului absolut cu bază mobilă, după relația:

Sporul absolut cu bază mobilă este:

$$\Delta_{4/3} = y_4 - y_3 = 11385 - 10367 = 1018 \text{ milioane USD}$$

$$\Delta_{i/i-1} = y_i - y_{i-1}$$

Interpretare:

Volumul exporturilor României a crescut, în anul 2001 față de anul 1997, cu 1018 milioane USD.

3. Volumul exporturilor României înregistrat în perioada 1997-2001 este prezentat în tabelul de mai jos:

Anii	1997	1998	1999	2000	2001
Valoare exporturi (milioane USD)	8431	8302	8487	10367	11385

Sursa : Anuarul statistic al României, C.N.S. 2002, p.449

Să se calculeze cu cât s-a diminuat, în procente, volumul exporturilor României în anul 1998 față de anul 1997.

Valoarea relativă a reducerii volumului exporturilor României, în anul 1998 față de anul 1997, se află prin calculul ritmului sporului cu bază fixă, după relația:

$$r_{i/0} = R_{i/0} - 100,$$

unde:

$$R_{i/0} = \frac{y_i}{y_0} \cdot 100$$

Ritmul sporului cu bază fixă este:

$$R_{1/0} = \frac{8302}{8431} \cdot 100 = 98,46\%$$

$$r_{1/0} = R_{1/0} - 100 = 98,46 - 100 = -1,54\%$$

Interpretare: Volumul exporturilor României s-a diminuat, în anul 1998 față de anul 1997, cu 1,54%.

4. Volumul exporturilor României înregistrat în perioada 1997-2001 este prezentat în tabelul de mai jos :

Anii	1997	1998	1999	2000	2001
Valoare exporturi (milioane USD)	8431	8302	8487	10367	11385

Sursa : Anuarul statistic al României, C.N.S. 2002, p. 449

Să se calculeze de câte ori a crescut volumul exporturilor României în anul 1999 față de anul 1997.

Valoarea creșterii volumului exporturilor României, în anul 1999 față de anul 1997, se află prin calculul ritmului variației cu bază fixă, după relația:

$$R_{i/0} = \frac{y_i}{y_0} \cdot 100$$

Ritmul variației cu bază fixă este:

$$R_{3/0} = \frac{8487}{8431} = 1,0066 \text{ ori}$$

Interpretare: Volumul exporturilor României a crescut, în anul 1998 față de anul 1997, de 1,0066 ori.

5. Volumul exporturilor României înregistrat în perioada 1997-2001 este prezentat în tabelul de mai jos :

Anii	1997	1998	1999	2000	2001
Valoare exporturi (milioane USD)	8431	8302	8487	10367	11385

Sursa : Anuarul statistic al României, C.N.S. 2002, p.449

Să se calculeze valoarea absolută a unui procent de reducere a volumului comerțului exterior al României înregistrată în anul 1998 față de anul 1997.

Valoarea absolută a unui procent de reducere a volumului comerțului exterior al României în anul 1998 față de anul 1997 se determină după relația:

$$(\%)y_{i/0} = \frac{\Delta_{i/0}}{r_{i/0}},$$

unde:

$\Delta_{i/0}$ este sporul absolut cu bază fixă, $\Delta_{i/0} = y_i - y_0$;

$r_{i/0}$ este ritmul sporului cu bază fixă, $r_{i/0} = R_{i/0} - 100$.

Sporul absolut cu bază fixă este:

$$\Delta_{1/0} = y_1 - y_0 = 8302 - 8431 = -129 \text{ mil.US}$$

Ritmul sporului cu bază fixă este:

$$r_{1/0} = R_{1/0} - 100 = 98,46 - 100 = -1,54\%$$

$$R_{1/0} = \frac{8302}{8431} \cdot 100 = 98,46\%$$

Valoarea absolută a unui procent de reducere a volumului comerțului exterior al României, în anul 1998 față de anul 1997, este:

$$(\%)y_{1/0} = \frac{\Delta_{1/0}}{r_{1/0}} = \frac{-129}{-1,54} = 83,76 \text{ milioane}$$

Interpretare: Valoarea absolută a unui procent de reducere a volumului exporturilor României înregistrată în anul 1998 față de anul 1997 este de 83,76 milioane USD.

6. Volumul exporturilor României înregistrat în perioada 1997-2001 este prezentat în tabelul de mai jos :

Anii	1997	1998	1999	2000	2001
Valoare exporturi (milioane USD)	8431	8302	8487	10367	11385

Sursa : Anuarul statistic al României, C.N.S. 2002, p.449

Să se calculeze valoarea absolută a unui procent de creștere a volumului exporturilor României înregistrată în anul 2001 față de anul 2000.

Valoarea absolută a unui procent de creștere a volumului exporturilor României în anul 2001 față de anul 2000 se calculează după relația :

$$(\%)y_{i/i-1} = \frac{\Delta_{i/i-1}}{r_{i/i-1}},$$

unde:

$\Delta_{i/i-1}$ este sporul absolut cu bază mobilă, $\Delta_{i/i-1} = y_i - y_{i-1}$;

$r_{i/i-1}$ este ritmul sporului cu bază mobilă, $r_{i/i-1} = R_{i/i-1} - 100$.

Sporul absolut cu bază mobilă este:

$$\Delta_{4/3} = y_4 - y_3 = 11385 - 10367 = 1018 \text{ milioane USD}$$

Ritmul sporului cu bază mobilă este:

$$r_{4/3} = R_{4/3} - 100 = 109,82 - 100 = 9,82\%$$

unde:

$$R_{4/3} = \frac{11385}{10367} \cdot 100 = 109,82\%$$

Valoarea absolută a unui procent de creștere a volumului exporturilor României în anul 2001 față de anul 2000 este:

$$(\%)y_{4/3} = \frac{\Delta_{4/3}}{r_{4/3}} = \frac{1018}{9,82} = 103,66 \text{ mil. USD}$$

Interpretare: Valoarea absolută a unui procent de creștere a volumului exporturilor României înregistrată în anul 2001 față de anul 2000 este de 103,66 milioane USD.

7. Volumul exporturilor României înregistrat în perioada 1997-2001 este prezentat în tabelul de mai jos:

Anii	1997	1998	1999	2000	2001
Valoare exporturi (milioane USD)	8431	8302	8487	10367	11385

Sursa : Anuarul statistic al României, C.N.S. 2002, p.449

Să se calculeze volumul mediu anual al exporturilor României în perioada 1997-2001.

Pentru o serie de timp de intervale, nivelul mediu se află calculând media aritmetică a termenilor seriei:

$$\bar{y} = \frac{y_1 + y_2 + \dots + y_{n-1} + y_n}{n}$$

Volumul mediu anual al exporturilor României în perioada 1997-2001 este:

$$\begin{aligned} \bar{y} &= \frac{8431 + 8302 + 8487 + 10367 + 11385}{5} \\ &= \frac{46972}{5} = 9394,4 \text{ milioane USD} \end{aligned}$$

Interpretare: Volumul mediu anual al exporturilor României în perioada 1997-2001 este de 9394,4 milioane USD.

8. Populația ocupată a României în perioada 1980-2001 este prezentată în tabelul de mai jos:

Anii	1980	1996	1999	2000	2001
Populația ocupată (mii persoane)	10350	9379	8420	8629	8563

Sursa : Anuarul statistic al României, 1982, DCS, p. 60,
Anuarul statistic al României, 2002, INS, p. 94

Să se calculeze numărul mediu anual al salariaților României din perioada 1980-2001.

Pentru o serie de timp de momente, cu momentele inegal distanțate, nivelul mediu se calculează ca o medie cronologică ponderată.

Numărul mediu anual al salariaților României, în perioada 1980-2001 este:

$$\bar{y}_{cr} = \frac{\frac{y_1+y_2}{2} \cdot t_1 + \frac{y_2+y_3}{2} \cdot t_2 + \dots + \frac{y_{n-1}+y_n}{2} \cdot t_{n-1}}{t_1+t_2+\dots+t_{n-1}}$$

$$\bar{y}_{cr} = \frac{\frac{10350+9379}{2} \cdot 16 + \frac{9379+8420}{2} \cdot 3}{21} + \frac{\frac{8420+8629}{2} \cdot 1 + \frac{8629+8563}{2} \cdot 1}{21} = 9602,4 \approx 9602 \text{ mii persoane}$$

Interpretare: Numărul mediu anual al populației ocupate în România, în perioada 1980-2001, este de 9602 mii persoane.

9. Efectivul angajaților unei firme la 1.01 este prezentat în tabelul de mai jos:

Anii	1992	1994	1996	1998	2000
Efectivul	96	103	300	304	302

Să se calculeze numărul mediu anual al angajaților firmei din perioada 1992-2000.

Pentru o serie de timp de momente, cu momentele egal distanțate, nivelul mediu se calculează ca o medie cronologică simplă.

$$\bar{y}_{cr} = \frac{\frac{y_1}{2} + y_2 + \dots + y_{n-1} + \frac{y_n}{2}}{n-1}$$

Numărul mediu anual al angajaților firmei, în perioada 1992-2000, este:

$$\bar{y}_{cr} = \frac{\frac{96}{2} + 103 + 300 + 304 + \frac{302}{2}}{5-1} = 226,5$$

Interpretare: Numărul mediu anual al efectivului firmei, în perioada 1992-2000, este de 226,5 ~ 227 persoane.

10. Ritmul mediu anual al variației volumului comerțului exterior al României din perioada 1990-2002 este de 0,9884. Se cere să se calculeze ritmul mediu anual al sporului.

Relație de calcul:

$$\bar{r} = \bar{R} - 1$$

$$\bar{r} = 0,9884 - 1 = -0,0116$$

Interpretare: Volumul comerțului exterior al României s-a redus în medie cu 1,16% pe an, în perioada 1990-2002.

11. Volumul producției și prețurile unitare pentru trei produse sunt prezentate în tabelul de mai jos:

Produsul	Volumul producției		Prețul unitar	
	2000	2002	2000	2002
A (m)	400	840	25	30
B (buc.)	500	600	12	10
C (kg)	600	960	8	10

Să se calculeze valoarea indicelui tip Laspeyres al volumului producției.

Indicele tip Laspeyres al volumului fizic al producției se determină după relația:

$$I_{1/0}^q = \frac{\sum_i p_{i0} q_{i1}}{\sum_i p_{i0} q_{i0}} \cdot 100$$

Elementele necesare determinării indicelui tip Laspeyres al volumului fizic al producției sunt prezentate în tabelul 8.1.

$$I_{1/0}^q = \frac{35880}{20800} \cdot 100 = 172,50\%$$

Interpretare: Volumul fizic al producției a crescut cu 72,5%, în anul 2002, față de anul 2000, în condițiile prețurilor din anul 2000.

Tabelul 8.1. Elemente de calcul

Produsul	P_{i0}	q_{i0}	q_{i1}	$P_{i0} \cdot q_{i0}$	$P_{i0} \cdot q_{i1}$
A (m)	25	400	840	10000	21000
B (buc.)	12	500	600	6000	72000
C (kg)	8	600	960	4800	7680
TOTAL	-	-	-	20800	35880

12. Volumul producției și prețurile unitare pentru trei produse sunt prezentate în tabelul de mai jos:

Produsul	Volumul producției		Prețul unitar	
	2000	2002	2000	2002
A (m)	400	840	25	30
B (buc.)	500	600	12	10
C (kg)	600	960	8	10

Să se calculeze indicele Paasche al volumului producției.

Indicele tip Paasche al volumului fizic al producției se determină după relația :

$$I_{1/0}^q = \frac{\sum_i p_{i1} q_{i1}}{\sum_i p_{i1} q_{i0}} \cdot 100$$

Elementele necesare calculului indicelui tip Paasche al volumului fizic al producției sunt prezentate în tabelul 8.2.

$$I_{1/0}^q = \frac{40800}{23000} \cdot 100 = 177,30\%$$

Interpretare: Volumul fizic al producției a crescut cu 77,3%, în anul 2002 față de anul 2000, în condițiile prețurilor din anul 2002.

Tabelul 8.2. Elemente de calcul

Produsul	p_{i1}	q_{i0}	q_{i1}	$p_{i1} \cdot q_{i0}$	$p_{i1} \cdot q_{i1}$
A (m)	30	400	840	12000	25200
B (buc.)	10	500	600	5000	6000
C (kg)	10	600	960	6000	9600
TOTAL	-	-	-	23000	40800

13. Volumul producției și prețurile unitare pentru trei produse sunt prezentate în tabelul de mai jos:

Produsul	Volumul producției		Prețul unitar	
	2000	2002	2000	2002
A (m)	400	840	25	30
B (buc.)	500	600	12	10
C (kg)	600	960	8	10

Să se calculeze indicele tip Sidgwick-Drobisch al prețului.

Indicele prețurilor tip Sidgwick-Drobisch se determină după relația:

Elementele necesare calculului indicelui tip Sidgwick-Drobisch al prețului sunt prezentate în tabelul 8.3.

$$I_{1/0}^p = \frac{\sum_i p_{i1} \cdot q_{i0} + \sum_i p_{i0} \cdot q_{i1}}{2 \cdot \sum_i p_{i0} \cdot q_{i0}} \cdot 100$$

$$I_{1/0}^p = \frac{23000 + 40800}{2 \cdot 20800} \cdot 100 = 112\%$$

Interpretare: Prețurile pe ansamblul celor trei produse au crescut cu 12%, în anul 2002 față de anul 2000.

Tabelul 8.3. Elemente de calcul

Produs	p_{i0}	p_{i1}	q_{i0}	q_{i1}	$p_{i1} \cdot q_{i0}$	$p_{i0} \cdot q_{i1}$	$p_{i0} \cdot q_{i0}$	$p_{i1} \cdot q_{i1}$
A (m)	25	30	400	840	12000	25200	10000	21000
B (buc.)	12	10	500	600	5000	6000	6000	72000
C (kg)	8	10	600	960	6000	9600	4800	7680
Total	-	-	-	-	23000	40800	20800	35880

14. Volumul producției și prețurile unitare pentru trei produse sunt prezentate în tabelul de mai jos:

Produsul	Volumul producției		Prețul unitar	
	2000	2002	2000	2002
A (m)	400	840	25	30
B (buc.)	500	600	12	10
C (kg)	600	960	8	10

Să se calculeze indicii tip Laspeyres al prețului.

Indicele tip Laspeyres al prețului se determină după relația

$$I_{1/0}^p = \frac{\sum_i p_{i1} q_{i0}}{\sum_i p_{i0} q_{i0}} \cdot 100$$

Elementele necesare determinării indicelui tip Laspeyres al prețului sunt prezentate în tabelul 8.4.

$$I_{1/0}^p = \frac{23000}{20800} \cdot 100 = 110,50\%$$

Interpretare: Prețurile pe ansamblul celor trei produse au crescut cu 10,5%, în anul 2002, față de anul 2000, în condițiile volumului producției din anul 2000.

Tabelul 8.4. Elemente de calcul

Produs	p_{i0}	p_{i1}	q_{i0}	$p_{i0} \cdot q_{i0}$	$p_{i1} \cdot q_{i0}$
A (m)	25	30	400	10000	12000
B (buc.)	12	10	500	6000	5000
C (kg)	8	10	600	4800	6000
TOTAL	-	-	-	20800	23000

15. Volumul producției și prețurile unitare pentru trei produse sunt

Produsul	Volumul producției		Prețul unitar	
	2000	2002	2000	2002
A (m)	400	840	25	30
B (buc.)	500	600	12	10
C (kg)	600	960	8	10

Să se calculeze indicii tip Paasche al prețului.

Indicele tip Paasche al prețului se determină după relația:

Elementele necesare determinării indicelui prețului tip Paasche sunt prezentate în tabelul 8.5.

$$I^p_{1/0} = \frac{\sum_i p_{i1} q_{i1}}{\sum_i p_{i0} q_{i1}} \cdot 100$$

$$I^p_{1/0} = \frac{40800}{35880} \cdot 100 = 113,70\%$$

Interpretare: Prețurile pe ansamblul celor trei produse au crescut cu 13,7%, în anul 2002, față de anul 200, în condițiile volumului producției din anul 2000.

Tabelul 8.5. Elemente de calcul

Produs	p_{i0}	p_{i1}	q_{i1}	$p_{i1} \cdot q_{i1}$	$p_{i0} \cdot q_{i1}$
A (m)	25	30	840	25200	21000
B (buc.)	12	10	600	6000	72000
C (kg)	8	10	960	9600	7680
TOTAL	-	-	-	40800	35880

16. Volumul producției și prețurile unitare pentru trei produse sunt prezentate în tabelul de mai jos:

Produsul	Volumul producției		Prețul unitar	
	2000	2002	2000	2002
A (m)	400	840	25	30
B (buc.)	500	600	12	10
C (kg)	600	960	8	10

Să se calculeze indicele tip Edgeworth al prețului.

Indicele tip Edgeworth al prețului se determină după relația:

$$I^p_{1/0} = \frac{\sum_i p_{i1} q_{i0} + \sum_i p_{i1} q_{i1}}{\sum_i p_{i0} q_{i0} + \sum_i p_{i0} q_{i1}} \cdot 100$$

Elementele necesare determinării indicelui tip Edgeworth al prețului sunt prezentate în tabelul 8.6.

$$I^p_{1/0} = \frac{23000+40800}{20800+35880} \cdot 100 = \frac{63800}{56680} \cdot 100 = 112,50\%$$

Interpretare: Prețurile pe ansamblul celor trei produse au crescut cu 12,5%, în anul 2002, față de anul 2000.

Tabelul 8.6. Elemente de calcul

Produs	p_{i0}	p_{i1}	q_{i0}	q_{i1}	$p_{i1} \cdot q_{i0}$	$p_{i1} \cdot q_{i1}$	$p_{i0} \cdot q_{i0}$	$p_{i0} \cdot q_{i1}$
A (m)	25	30	400	840	12000	25200	10000	21000
B (buc.)	12	10	500	600	5000	6000	6000	72000
C (kg)	8	10	600	960	6000	9600	4800	7680
Total	-	-	-	-	23000	40800	20800	35880

17. Volumul producției și prețurile unitare pentru trei produse sunt prezentate în tabelul de mai jos:

Produsul	Volumul producției		Prețul unitar	
	2000	2002	2000	2002
A (m)	400	840	25	30
B (buc.)	500	600	12	10
C (kg)	600	960	8	10

Să se calculeze indicele mediu aritmetic.

Indicele mediu aritmetic al volumului fizic al producției se determină după relația:

$$I_{1/0}^q = \frac{\sum_i i_{1/0} \cdot y_{i0} \cdot p_{i0}}{\sum_i y_{i0} \cdot p_{i0}}$$

Elementele necesare determinării indicelui mediu aritmetic sunt prezentate în tabelul 8.7.

$$\begin{aligned} I_{1/0}^q &= \frac{10000 \cdot 2,1 + 6000 \cdot 1,2 + 4800 \cdot 1,6}{20800} = \\ &= \frac{35880}{20800} = 1,725 \end{aligned}$$

Interpretare: Volumul fizic al producției a crescut anual, în medie, cu 72,5%.

Tabelul 8.7. Elemente de calcul

Produsul	Valoarea producției obținute în perioada de bază	Indici individuali ai volumului fizic
A (m)	10000	2,1
B (buc.)	6000	1,2
C (kg)	4800	1,6
TOTAL	20800	-

18. Volumul producției și prețurile unitare pentru trei produse sunt prezentate în tabelul de mai jos:

Produsul	Volumul producției		Prețul unitar	
	2000	2002	2000	2002
A (m)	400	840	25	30
B (buc.)	500	600	12	10
C (kg)	600	960	8	10

Să se calculeze indicele mediu geometric.

Indicele mediu geometric – Irving Fischer se determină după relația:

$$I_{1/0} = \sqrt{\frac{\sum_i q_{i1} \cdot p_{i0}}{\sum_i q_{i0} \cdot p_{i0}} \cdot \frac{\sum_i q_{i1} \cdot p_{i1}}{\sum_i q_{i0} \cdot p_{i1}}}$$

Elementele necesare determinării indicelui mediu geometric sunt prezentate în tabelul 8.8.

$$\begin{aligned} I_{1/0} &= \sqrt{\frac{35880}{20800} \cdot \frac{40800}{23000}} = \\ &= \sqrt{1,725 \cdot 1,773} = \sqrt{3,058} = \\ &= 1,748 \end{aligned}$$

Interpretare: Volumul fizic al producției a crescut anual, în medie, cu 74,8%.

Tabelul 8.8. Elemente de calcul

Produs	p_{i0}	p_{i1}	q_{i0}	q_{i1}	$p_{i1} \cdot q_{i0}$	$p_{i1} \cdot q_{i1}$
A (m)	25	30	400	840	12000	25200
B (buc.)	12	10	500	600	5000	6000
C (kg)	8	10	600	960	6000	9600
TOTAL	-	-	-	-	23000	40800

Produs	$p_{i0} \cdot q_{i0}$	$p_{i0} \cdot q_{i1}$
A (m)	10000	21000
B (buc.)	6000	72000
C (kg)	4800	7680
TOTAL	20800	35880

Capitolul 9

EȘANTIONAREA

- ◆ Ce este cercetarea prin sondaj ?
- ◆ Noțiuni și termeni perechi
- ◆ Demersul cercetării prin sondaj
- ◆ Probleme ale eșantionării
- ◆ Erori de reprezentativitate
- ◆ Calculul mărimii erorilor de reprezentativitate
- ◆ Alegeți răspunsul corect
- ◆ Răspunsuri corecte
- ◆ Probleme

9.1. Ce este cercetarea prin sondaj ?

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 270-273)

Cercetarea prin sondaj presupune cunoașterea unei populații pe baza datelor înregistrate numai asupra unei părți reprezentative, numită eșantion, extrasă din populație.

Domenii de aplicabilitate: agricultură, industrie, comerț, marketing, domeniul social, medicină etc.

Avantaje: cost redus, rapiditatea obținerii rezultatelor, cunoașterea unor colectivități greu accesibile sau perisabile etc.

Limite: apariția erorilor de înregistrare și a erorilor de reprezentativitate.

9.2. Noțiuni și termeni perechi

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 273-275)

- **Noțiuni**
 - *Colectivitate totală*. Definirea populației (colectivității totale) se face prin enumerarea tuturor unităților componente. Ansamblul acestora formează *volumul colectivității* și se notează cu N .
 - *Eșantion*. Reprezintă un subansamblu de unități extrase, după principiile selecției, dintr-o colectivitate pe care dorim să o cunoaștem. Volumul eșantionului se notează cu n . În cazul cercetării prin sondaj are loc relația: $n < N$. Când $n = N$, este cazul unei observări exhaustive.
 - *Variabila sau caracteristica de distribuție*. Este variabila după care sunt observate unitățile eșantionului, respectiv ale populației. Variabila observată poate fi cantitativă (venitul indivizilor, numărul angajaților unei firme) sau calitativă (sexul individului, tipul de produse). Valorile unei caracteristici X sunt x_i .

• **Termeni perechi : Parametri – statistici – valori tipice de sondaj**

- *Parametri – valori tipice ale distribuției populației.* Un parametru reprezintă o mărime fixă, reală, dar necunoscută a unei populații.

- *Statistici – variabile aleatoare de selecție sunt construite pe baza distribuțiilor de selecție.* Valorile unei variabile aleatoare de selecție reprezintă valorile posibile ale variabilei X în cele k eșantioane posibil de extras din populația de referință. Variabilele $X_1, X_2, \dots, X_n$ sunt independente stochastic în ansamblu și urmează aceeași lege de probabilitate cu variabila studiată X .

Statisticile sunt folosite pentru estimarea parametrilor unei populații și în acest caz se numesc *estimatori*, fiind notați cu semnul „ $\hat{\cdot}$ ” deasupra parametrului considerat. Astfel, pentru parametrul θ , estimatorul său poate fi notat $\hat{\theta}$.

- *Valorile tipice de sondaj – valori tipice ale distribuției eșantionului.* Sunt valori obținute prin prelucrarea datelor înregistrate pe un eșantion observat (de exemplu, media și varianța unui eșantion etc.).

În procesul estimării parametrilor unei populații, valorile tipice de sondaj se numesc *estimații*.

Tabelul 9.1 Termeni folosiți în procesul de estimare statistică

	Parametri (valori)	Estimatori (variabile)	Estimații (valori)
1. Media	μ	$\hat{\mu}$	$\bar{x}$
2. Varianța	σ^2	$\hat{\sigma}^2$	s^2, s'^2
3. Proporția	p	$\hat{p}$	f
4. Coeficienții de regresie	α, β	$\hat{\alpha}, \hat{\beta}$	a, b
5. Coeficientul de corelație	ρ	$\hat{\rho}$	r_y

9.3. Demersul cercetării prin sondaj

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 270-273)

Demersul cercetării prin sondaj presupune parcurgerea unui întreg proces, de la culegerea datelor pe baza eșantionului, până la extinderea și analiza rezultatelor la nivelul colectivității totale, în condiții de cost și precizie admisibile.

9.4. Probleme ale eșantionării

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 278-285)

- **Definirea populației**
 - Delimitarea populației;
 - Verificarea gradului de omogenitate.
- **Alegerea bazei de sondaj** (lista sau câmpul de anchetă). Baza de sondaj reprezintă populația pentru care se extind rezultatele obținute prin sondaj.
- **Definirea eșantionului**
 - Alegerea și definirea unităților folosite în eșantionare: unitatea de observare și unitatea de eșantionare;
 - Alegerea eșantionului: principii de eșantionare; alegerea tipului de sondaj; alegerea procedurilor de extragere a unităților de eșantionare; alegerea gradului de reprezentativitate și a mărimii eșantionului.

Observație ! Este esențial să se aleagă cel mai potrivit tip de sondaj pentru populația observată, adică un sondaj care să dea posibilitatea obținerii unui eșantion reprezentativ, fără „biais“ (deplasare) față de elementele populației din care este extras eșantionul. În caz contrar, un eșantion nereprezentativ ar da informații care nu ar estima corect populația și, logic, deciziile fundamentate pe o astfel de informație ar fi greșite.

9.5. Erori de reprezentativitate

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 287-288)

- **Erori de reprezentativitate**
 - Erori sistematice de reprezentativitate
 - Erori întâmplătoare de reprezentativitate

Observație ! Pe baza erorilor de reprezentativitate se poate stabili precizia estimațiilor parametrilor colectivității totale.

- **Erori de acoperire**

Erorile de acoperire sunt datorate bazei incomplete de sondaj și non-răspunsurilor unor indivizi la întrebările puse printr-un chestionar. Acestea duc la neacoperirea în întregime a populației pe care vrem să o studiem și au ca efect un *biais al estimatorilor* ce nu poate fi măsurat (astfel de erori trebuie eliminate printr-o mai bună organizare a cercetării prin sondaj).

9.6. Calculul mărimii erorilor de reprezentativitate

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 288-290)

- Deplasarea sau *biais-ul*
- Eroarea medie de reprezentativitate
- Eroarea limită

Toți acești indicatori sunt mărimi ale variației. Folosirea lor în calitate de măsuri ale erorilor de reprezentativitate este posibilă deoarece acestea *exprimă abateri ale valorilor tipice de sondaj față de parametrii colectivității generale*, iar abaterile – gradul de împrăștiere – se măsoară prin indicatori ai dispersiei.



9.7. Alegeți răspunsul corect

1. Într-o cercetare prin sondaj, valoarea medie la nivelul unei populații reprezintă
 - a) un parametru
 - b) o valoare tipică de sondaj
 - c) un estimator
 - d) o estimație
2. Valoarea medie la nivelul unui eșantion, într-o cercetare prin sondaj, reprezintă
 - a) un parametru
 - b) o valoare tipică de sondaj
 - c) un estimator
 - d) o estimație
3. Formarea eșantionului în cazul unui sondaj aleator simplu presupune ca
 - a) fiecărui individ din populație să i se asocieze o probabilitate, egală sau inegală, de a intra în eșantion
 - b) eșantionul să se formeze prin extragerea unui număr de grappe
 - c) eșantionul să se formeze prin extragerea aleatoare a unităților la nivelul fiecărei grupe
4. În cazul sondajului stratificat (tipic) simplu, eșantionul se formează extrăgând din fiecare grupă omogenă a colectivității generale
 - a) același număr de unități, indiferent de ponderea fiecărei grupe în colectivitate
 - b) un număr de unități proporțional cu ponderea grupelor în colectivitatea generală
 - c) un număr de unități proporțional cu ponderea grupelor în colectivitatea generală și gradul de dispersie din grupe
5. În cazul sondajului stratificat (tipic) proporțional, eșantionul se formează extrăgând din fiecare grupă omogenă a colectivității generale
 - a) același număr de unități, indiferent de ponderea fiecărei grupe în colectivitate
 - b) un număr de unități proporțional cu ponderea grupelor în colectivitatea generală
 - c) un număr de unități proporțional cu ponderea grupelor în colectivitatea generală și gradul de dispersie din grupe

6. În cazul sondajului stratificat (tipic) optim, eșantionul se formează extrăgând din fiecare grupă omogenă a colectivității generale
- a) același număr de unități, indiferent de ponderea fiecărei grupe în colectivitate
 - b) un număr de unități proporțional cu ponderea grupelor în colectivitatea generală
 - c) un număr de unități proporțional cu ponderea grupelor în colectivitatea generală și gradul de dispersie din grupe
7. În cazul unei cercetări prin sondaj se pot produce erori de
- a) reprezentativitate
 - b) respingere pe nedrept a unei ipoteze adevărate
 - c) acoperire a populației
8. Erorile de reprezentativitate reprezintă diferențe între
- a) valoarea estimată și valoarea adevărată a unui parametru
 - b) valorile reale și valorile înregistrate
 - c) populația reală și eșantionul observat
9. După sursa de proveniență, erorile de reprezentativitate pot fi
- a) erori sistematice
 - b) erori întâmplătoare
 - c) erori de acoperire
10. Sunt erori probabile de reprezentativitate
- a) eroarea medie de reprezentativitate
 - b) eroarea sistematică
 - c) eroarea limită
11. Mărimea erorii limită depinde de
- a) mărimea erorii medii de reprezentativitate
 - b) volumul eșantionului
 - c) coeficientul de probabilitate Z

$$n = \frac{1,96^2 \cdot 21}{4} = 2017 \text{ persoane}$$

Interpretare: Numărul persoanelor care ar trebui extrase aleatoriu simple repetat din populația de volum $N=200$ este de 2017-20 persoane.



9.8. Răspunsuri corecte

1. a) Într-o cercetare prin sondaj, valoarea medie la nivelul unei populații reprezintă un parametru.
2. b) și d) Valoarea medie la nivelul unei eșantion, într-o cercetare prin sondaj, reprezintă o valoare tipică de sondaj, respectiv, o estimatie.
3. a) Sondajul aleator simplu presupune ca fiecărui individ din populație să i se asocieze o probabilitate, egală sau inegală, de a intra în eșantion.
4. a) În cazul sondajului stratificat (tipic) simplu, eșantionul se formează extrăgând din fiecare grupă omogenă a colectivității generale același număr de unități, indiferent de ponderea fiecărei grupe în colectivitate.
5. b) În cazul sondajului stratificat (tipic) proporțional, eșantionul se formează extrăgând din fiecare grupă omogenă a colectivității generale un număr de unități proporțional cu ponderea grupelor în colectivitatea generală.
6. c) În cazul sondajului stratificat (tipic) optim, eșantionul se formează extrăgând din fiecare grupă omogenă a colectivității generale un număr de unități proporțional cu ponderea grupelor în colectivitatea generală și gradul de dispersie din grupe.
7. a) și c) În cazul unei cercetări prin sondaj se pot produce erori de reprezentativitate și de acoperire a populației.
8. a) Erorile de reprezentativitate sunt diferențe între valoarea estimată și valoarea adevărată a unui parametru.
9. a) și b) După sursa de proveniență, erorile de reprezentativitate pot fi erori sistematice și erori întâmplătoare.
10. a) și c) Sunt erori probabile de reprezentativitate: eroarea medie de reprezentativitate și eroarea limită.
11. a), b) și c) Mărimea erorii limită depinde de mărimea erorii medii de reprezentativitate, volumul eșantionului și coeficientul de probabilitate Z .



9.9. Probleme

9.9.1. Determinarea volumului eșantionului

1. Dispunem de următoarele date la nivelul unei populații: $N=400$ persoane, $\sigma^2=100$. Admitem o precizie de $\Delta_{\hat{\mu}}=2$ și un risc $\alpha=0,05$, pentru care $z=1,96$. Se cere să se calculeze volumul eșantionului corespunzător unui sondaj aleator simplu nerepetat. Interpretați rezultatul.

Volumul eșantionului corespunzător unui sondaj aleator simplu nerepetat se calculează după relația:

$$n = \frac{z^2 \cdot \sigma^2}{\Delta_{\hat{\mu}}^2 + \frac{z^2 \cdot \sigma^2}{N}}$$

Volumul eșantionului este:

$$n = \frac{1,96^2 \cdot 100}{4 + \frac{1,96^2 \cdot 100}{400}} = \frac{384,16}{4,96} = 77,45 \text{ pers.}$$

Interpretare: Numărul persoanelor care ar trebui extrase aleatoriu simplu și nerepetat din populația de volum $N=400$ este de 77,45~77 persoane.

2. Cunoaștem următoarele date la nivelul unei populații: $N=200$ persoane, $\sigma^2=21$. Admitem o precizie de $\Delta_{\hat{\mu}}=2$ și un risc $\alpha=0,05$, pentru care $z=1,96$. Se cere să se calculeze volumul eșantionului corespunzător unui sondaj aleatoriu simplu repetat. Interpretați rezultatul.

Volumul eșantionului corespunzător unui sondaj aleator simplu repetat se calculează după relația:

$$n = \frac{z^2 \cdot \sigma^2}{\Delta_{\hat{\mu}}^2}$$

Volumul eșantionului este:

$$n = \frac{1,96^2 \cdot 21}{4} = 20,17 \text{ persoane}$$

Interpretare: Numărul persoanelor care ar trebui extrase aleatoriu simplu repetat din populația de volum $N=200$ este de 20,17~20 persoane.

9.9.2. Repartizarea volumului eșantionului pe grupe tipice

1. Pentru o populație structurată pe 2 grupe: $N_1=600$ bărbați și $N_2=400$ femei, se organizează un sondaj aleator stratificat simplu și se extrage un eșantion de volum $n=100$. Se cere să se repartizeze volumul eșantionului pe grupe, corespunzător acestui tip de sondaj.

În cazul unui sondaj aleator stratificat simplu, repartizarea pe grupe a numărului de unități la nivelul eșantionului se efectuează după relația:

$$n_i = \frac{n}{k},$$

unde:

$k=2$ grupe.

Numărul de unități pe grupe la nivelul eșantionului este:

$$n_i = \frac{n}{k} = \frac{100}{2} = 50.$$

Interpretare: În condițiile organizării unui sondaj stratificat simplu, volumul eșantionului ($n=100$ unități) se repartizează în mod egal pe grupe. Din fiecare grupă a populației, indiferent de ponderea acesteia în totalul populației, vor fi extrase 50 persoane.

2. Pentru o populație structurată pe 2 grupe: $N_1=600$ bărbați și $N_2=400$ femei, se organizează un sondaj aleator stratificat proporțional și se extrage un eșantion de volum $n=100$. Se cere să se repartizeze volumul eșantionului pe grupe, corespunzător acestui tip de sondaj.

În cazul unui sondaj aleator stratificat proporțional, repartizarea pe grupe a numărului de unități la nivelul eșantionului se efectuează după relația:

$$n_i = \frac{N_i}{N} \cdot n$$

Numărul de unități pe grupe la nivelul eșantionului este:

$$n_1 = \frac{N_1}{N} \cdot n = \frac{600}{1000} \cdot 100 = 60 \text{ bărbați}$$

$$n_2 = \frac{N_2}{N} \cdot n = \frac{400}{1000} \cdot 100 = 40 \text{ femei}$$

Interpretare: În condițiile organizării unui sondaj stratificat proporțional, volumul eșantionului se repartizează pe grupe ale populației, proporțional cu ponderea

acestora în totalul populației. Din prima grupă a populației vor fi extrași 60 bărbați, iar din a doua grupă vor fi extrase 40 femei.

3. Pentru o populație structurată pe 2 grupe: $N_1=600$ bărbați și $N_2=400$ femei, $\sigma_1^2 = 25$, $\sigma_2^2 = 36$, se organizează un sondaj aleator stratificat optim și se extrage un eșantion de volum $n=100$. Se cere să se repartizeze volumul eșantionului pe grupe, corespunzător acestui tip de sondaj.

În cazul unui sondaj aleator stratificat optim, repartizarea pe grupe a numărului de unități la nivelul eșantionului se efectuează după relația:

$$n_i = \frac{N_i \cdot \sigma_i}{\sum_i N_i \cdot \sigma_i} \cdot n$$

Numărul de unități pe grupe la nivelul eșantionului este:

$$n_1 = \frac{N_1 \cdot \sigma_1}{\sum_i N_i \cdot \sigma_i} \cdot n = \frac{500 \cdot 5}{500 \cdot 5 + 400 \cdot 6} \cdot 100 = 51,02 \approx 51$$

$$n_2 = \frac{N_2 \cdot \sigma_2}{\sum_i N_i \cdot \sigma_i} \cdot n = \frac{400 \cdot 6}{500 \cdot 5 + 400 \cdot 6} \cdot 100 = 48,98 \approx 49$$

Interpretare: În condițiile organizării unui sondaj stratificat optim, volumul eșantionului format din 100 unități se repartizează pe grupe astfel: din prima grupă vor fi extrași 51 bărbați, iar din a doua grupă vor fi extrase 49 femei.

9.9.3. Determinarea numărului de eșantioane posibil de extras dintr-o populație

1. Să se afle câte eșantioane de volum $n=3$ se pot extrage aleator simplu repetat dintr-o populație de volum $N=5$.

Numărul de eșantioane care se pot extrage aleatoriu simplu repetat dintr-o populație se calculează după relația:

$$k = N^n$$

Numărul de eșantioane este:

$$k = 5^3 = 125 \text{ eșantioane}$$

Interpretare: Din populația considerată ($N=5$ unități) se pot extrage aleator simplu repetat un număr de 125 eșantioane de volum $n=3$ unități.

2. Să se afle câte eșantioane de volum $n=3$ se pot extrage aleator simplu nerepetat dintr-o populație de volum $N=5$.

Numărul de eșantioane care se pot extrage aleatoriu simplu nerepetat dintr-o populație se calculează după relația:

$$k = C_N^n = \frac{N!}{(N-n)!n!}$$

Numărul de eșantioane este:

$$k = C_5^3 = \frac{5!}{(5-3)!3!} = 10 \text{ eșantioane.}$$

Interpretare: Din populația considerată ($N=5$ unități) se pot extrage aleator simplu nerepetat, un număr de 10 eșantioane de volum $n=3$ unități.

Capitolul 10

ESTIMAREA PARAMETRILOR UNEI POPULAȚII

- ◆ Probleme generale ale unei estimări
- ◆ Proprietăți ale estimatorilor
- ◆ Estimarea punctuală a parametrilor unei populații
- ◆ Estimarea prin interval de încredere a parametrilor unei populații
- ◆ Determinarea volumului eșantionului
- ◆ Estimarea parametrilor unei populații cu SPSS
- ◆ Alegeți răspunsul corect
- ◆ Recunoașteți exemplul corect
- ◆ Răspunsuri corecte
- ◆ Probleme

10.1. Probleme generale ale unei estimări

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 291-296)

Contribuții la teoria estimației: A.M. Legendre (1805) și C.F. Gauss (1809), P.S. Laplace și Gauss, K. Pearson, R.A. Fisher („Teoria statistică a estimației“ (1925)).

Noțiuni

- **Estimare.** Estimarea este un procedeu prin care se află valoarea unui parametru al unei populații pe baza datelor înregistrate la nivelul unui eșantion extras din aceasta. Estimarea se poate efectua fie sub formă de estimare punctuală, fie sub formă de estimare prin interval de încredere.

- Estimarea punctuală

- Estimarea prin interval de încredere

- **Estimatori.** Un estimator este o statistică (o variabilă aleatoare) care urmează o lege de probabilitate care depinde, în general, de un parametru necunoscut. Un estimator este utilizat pentru a estima un parametru al unei populații.

Pentru un parametru θ , un estimator se notează cu $\hat{\theta}$ și reprezintă o funcție de n variabile aleatoare $(X_1, X_2, \dots, X_n)$ independente și identic distribuite: $\hat{\theta} = f(X_1, X_2, \dots, X_n)$.

- **Estimații.** O estimație reprezintă o valoare $\hat{\theta}_i$ a unui estimator $\hat{\theta}$ al parametrului θ , calculată pe baza unui eșantion n valori $(x_1, x_2, \dots, x_n)$, prelevat dintr-o populație N , adică este o valoare tipică de sondaj de forma: $f(x_1, x_2, \dots, x_n)$.

- **Distribuție de selecție.** O distribuție de selecție este distribuția unei statistici. Dintr-o populație de volum N , de parametri μ (media populației) și σ^2 (varianța sa), se pot extrage k eșantioane de volum n . Fiecare eșantion are media $(\bar{x}_j)$, proporția (f_j) și varianța sa (s_j^2) . Pe ansamblul celor k eșantioane, se obțin variabilele:

$\hat{\mu}: (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k)$;

$$\hat{p} : (f_1, f_2, \dots, f_k);$$

$$\hat{\sigma}^2 : (s_1^2, s_2^2, \dots, s_k^2),$$

unde:

$\hat{\mu}$, medie de selecție;

$\hat{p}$, proporție de selecție;

$\hat{\sigma}^2$, varianță de selecție.

$\hat{\mu}$, $\hat{p}$, $\hat{\sigma}^2$ apar ca variabile, cu niveluri diferite pentru fiecare eșantion. Pentru fiecare variabilă se determină un nivel mediu $M(\hat{\theta})$ și o varianță $V(\hat{\theta})$ notată și $\sigma_{\hat{\theta}}^2$.

10.2. Proprietăți ale estimatorilor

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 297-298)

- Nedeplasare
- Convergență
- Eficiență

10.3. Estimarea punctuală a parametrilor unei populații

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 299-307)

- Estimarea punctuală a parametrilor unei populații presupune estimarea unei valori posibile a estimatorului parametrului căutat, pe baza datelor înregistrate la nivelul unui eșantion.

10.3.1. Estimatorul $\hat{\mu}$ al mediei μ a unei populații

Media μ a populației se poate estima punctual prin media ($\bar{x}$), obținută la nivelul unui eșantion. Media $\bar{x}$ este o valoare a estimatorului $\hat{\mu}$, calculată pe baza datelor de la nivelul unui eșantion.

- **Proprietățile estimatorului $\hat{\mu}$.** Estimatorul mediei urmează legea de probabilitate normală. Este un estimator *nedeplasat, convergent și eficient*.

Distribuția mediei de selecție se fundamentează pe **teorema limită centrală (TLC)**: suma unui număr suficient de mare de variabile (X) aleatoare independente și identic distribuite urmează aproximativ o lege normală.

Distribuția mediei de selecție tinde spre o distribuție normală, atunci când:

- volumul eșantionului, n , tinde spre infinit, indiferent de legea de distribuție urmată de variabila aleatoare de distribuție a populației de origine;
- volumul eșantionului este oricât de mic, dacă variabila aleatoare de distribuție a populației urmează o distribuție aproximativ normală.

- **Caracteristici ale distribuției mediei de selecție**

- $\hat{\mu}$ urmează întotdeauna o lege normală sau aproximativ normală, de medie μ și varianță $\sigma_{\hat{\mu}}^2$, respectiv: $\hat{\mu} \sim N(\mu, \sigma_{\hat{\mu}}^2)$;
- Media distribuției mediei de selecție este egală cu media populației: $M(\hat{\mu}) = \mu$;
- Varianța mediei de selecție ($\sigma_{\hat{\mu}}^2$) este egală cu varianța populației (σ^2) împărțită la volumul eșantionului (n):

1. pentru o populație infinită (cazul sondajului aleator repetat),

$$\sigma_{\hat{\mu}}^2 = \frac{\sigma^2}{n};$$

2. pentru o populație finită (cazul sondajului aleator nerepetat),

$$\sigma_{\hat{\mu}}^2 = \frac{\sigma^2}{n} \frac{N-n}{N},$$

unde: σ^2 este varianța populației;

n – volumul eșantionului;

N – volumul populației.

Observație! Variația mediei de selecție stă la baza calculului erorii medii de reprezentativitate, care se măsoară în unități de abateri standard, adică prin abaterea medie pătratică a mediei de selecție față de media ei ($\sigma_{\bar{x}}$).

10.3.2. Estimatorul $\hat{p}$ al proporției p a unei populații

Estimatorul proporției p se notează cu $\hat{p}$. Este o variabilă aleatoare care urmează o lege Bernoulli, de medie $p = \frac{N_A}{N}$,

unde: N_A – indivizii din categoria A;

N – întreaga populație.

O valoare posibilă a estimatorului $\hat{p}$ este proporția (f) calculată la nivelul unui eșantion. Proporția f este o *estimație* definită prin:

$$f = \frac{n_A}{n},$$

unde n_A este numărul unităților din eșantion care posedă atributul considerat, iar n , volumul eșantionului.

- **Proprietățile estimatorului proporției $\hat{p}$.** Este un *estimator nedeplasat și convergent* al proporției p .

10.3.3. Estimatorul $\hat{\sigma}^2$ al varianței σ^2 a unei populații

Estimarea punctuală a varianței σ^2 presupune calculul unei estimații s^2 pe baza datelor unui eșantion prelevat din populație.

O estimație calculată la nivelul unui eșantion de volum n se poate afla după relația: $s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$. O estimație nedeplasată a varianței mediei de selecție se determină prin relația:

$$s_{\hat{\mu}}^2 = \frac{s'^2}{n}, \text{ unde } s'^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}.$$

- **Proprietățile estimatorului $\hat{\sigma}^2$.** Estimatorul $\hat{\sigma}^2$ este un *estimator deplasat*.

10.3.4. Estimatorul varianței distribuției de selecție a diferenței dintre două medii

- **Varianța de selecție a diferenței dintre două medii se calculează după relația:**

$$\sigma_{\hat{\mu}_1 - \hat{\mu}_2}^2 = \sigma_{\hat{\mu}_1}^2 + \sigma_{\hat{\mu}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

– Când varianțele celor două populații comparate sunt diferite între ele, $\sigma_1^2 \neq \sigma_2^2$. Estimatorul varianței de selecție a diferenței dintre două medii este:

$$\hat{\sigma}_{\hat{\mu}_1 - \hat{\mu}_2}^2 = \frac{\hat{\sigma}_1'^2}{n_1} + \frac{\hat{\sigma}_2'^2}{n_2}$$

– Când varianțele celor două populații comparate sunt egale între ele, $(\sigma_1^2 = \sigma_2^2 = \sigma^2)$. Estimatorul varianței de selecție a diferenței dintre două medii este:

$$\hat{\sigma}_{\hat{\mu}_1 - \hat{\mu}_2}^2 = \hat{\sigma}_w^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right), \text{ unde } \hat{\sigma}_w^2 = \frac{\hat{\sigma}_1'^2(n_1 - 1) + \hat{\sigma}_2'^2(n_2 - 1)}{n_1 + n_2 - 2}.$$

Estimația varianței de selecție a diferenței dintre două medii se calculează după relația:

$$s_{\hat{\mu}_1 - \hat{\mu}_2}^2 = s_w^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right), \text{ unde } s_w^2 = \frac{s_1'^2(n_1 - 1) + s_2'^2(n_2 - 1)}{n_1 + n_2 - 2}.$$

s_w^2 reprezintă **varianța ponderată** calculată pe baza estimațiilor varianțelor calculate la nivelul eșantioanelor n_1, n_2 .

10.3.5. Estimatorul varianței distribuției de selecție a diferenței dintre două proporții

Varianța de selecție a diferenței dintre două proporții se calculează în aceeași manieră ca varianța diferenței dintre două medii, după relația:

$$\sigma_{\hat{p}_1 - \hat{p}_2}^2 = \sigma_{\hat{p}_1}^2 + \sigma_{\hat{p}_2}^2 = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2},$$

unde p_1 și p_2 sunt proporțiile populațiilor comparate.

$$\text{Estimatorul său este: } \hat{\sigma}_{\hat{p}_1 - \hat{p}_2}^2 = \hat{\sigma}_{\hat{p}_1}^2 + \hat{\sigma}_{\hat{p}_2}^2 = \frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}.$$

10.4. Estimarea prin interval de încredere a parametrilor unei populații

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economică, București, 2002, pp. 307-313)

- Estimarea prin interval de încredere a parametrilor unei populații presupune **afllarea limitelor de încredere ale unui interval care acoperă valoarea adevărată a unui parametru al populației**. Estimarea prin interval de încredere ține seama de fluctuațiile distribuției de selecție a estimatorului parametrului considerat.

10.4.1. Definirea intervalului de încredere (I.C.)

A defini un interval de încredere înseamnă a căuta limitele de încredere, $L_i = \hat{\theta} - \Delta_{\hat{\theta}}$ și $L_s = \hat{\theta} + \Delta_{\hat{\theta}}$, care acoperă valoarea parametrului θ , pentru un coeficient de încredere: $P(L_i \leq \theta \leq L_s) = 1 - \alpha$, adică:

$$I.C. = [\hat{\theta} - \Delta_{\hat{\theta}}; \hat{\theta} + \Delta_{\hat{\theta}}],$$

unde:

L_i, L_s – limitele de încredere: inferioară, respectiv superioară;

$(1 - \alpha)$ – probabilitatea cu care se garantează că intervalul acoperă valoarea adevărată a parametrului θ , în cazul unei probleme de estimare, respectiv, valoarea unei statistici, în cazul unei probleme de distribuție de selecție;
 α – riscul, respectiv, probabilitatea că intervalul nu conține valoarea căutată.

Calculul limitelor intervalului de încredere pornește de la **estimarea valorii erorii limită**, $\Delta_{\hat{\theta}}$, pe baza distribuției de selecție a estimatorului $\hat{\theta}$ al parametrului θ . Determinarea erorii limită este necesară pentru realizarea unei estimări prin interval de încredere.

10.4.2. Estimarea mediei prin interval de încredere

Dacă parametrul căutat ar fi μ – media unei populații, iar $\hat{\mu}$ – media de selecție, construirea I.C. pleacă de la o ipoteză asupra distribuției mediilor de selecție, deci și a abaterilor medii pătratice ale acestora, față de media populației, respectiv, față de media lor.

Estimarea mediei prin I.C., în cazul în care se cunoaște legea populației și aceasta este o lege normală $N(\mu, \sigma^2)$, pot apare două situații: *când se cunoaște varianța și când nu se cunoaște varianța*.

- **Când se cunoaște varianța unei populații**, construirea intervalului de încredere se bazează pe variabila normală centrată redusă Z . Intervalul de încredere construit pe baza datelor unui eșantion este:

$$\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}},$$

unde:

$z_{\alpha/2}$ este o valoare a variabilei normale centrate reduse Z ,
 α este un nivel al probabilității, cuprins între zero și unu.

- **Când varianța este necunoscută**, construirea intervalului de încredere se bazează pe variabila t Student.

Intervalul de încredere construit pe baza datelor unui eșantion este:

$$\left(\bar{x} \pm t_{\alpha/2} \frac{s'}{\sqrt{n}} \right).$$

10.4.3. Estimarea proporției prin interval de încredere

Intervalul de încredere pentru parametrul p, construit pe baza datelor înregistrate la nivelul unui eșantion, este:

$$\left[f \pm z_{\alpha/2} \cdot \sqrt{\frac{f(1-f)}{n}} \right]$$

10.5. Estimarea parametrilor unei populații cu SPSS

(Vezi Elisabeta Jaba și Ana Grama, *Analiza statistică cu SPSS sub Windows*, Editura Polirom, Iași, 2004, pp. 193-196)

10.5.1. Estimarea mediei μ a unei populații

SPSS calculează valoarea tipică de sondaj, în acest caz, media eșantionului ($\bar{x}$). Pentru estimarea mediei μ a unei populații, folosim din Outputul „*Descriptives*“, obținut prin demersul: *Analyse*→*Descriptives Statistics*→*Explore*, următoarele rezultate: media eșantionului ($\bar{x}$), gradul de încredere corespunzător, eroarea standard a mediei ($\sigma_{\bar{\mu}}$), precum și limita inferioară și limita superioară ale intervalului de încredere.

Rezultatele afișate în fereastra Output „*Descriptives*“ sunt prezentate astfel:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
VAR00001	5	100,0%	0	,0%	5	100,0%

Descriptives

			Statistic	Std. Error
VAR00001	Mean		4,0000	,7071
	95% Confidence Interval for Mean	Lower Bound	2,0368	
		Upper Bound	5,9632	

	5% Trimmed Mean		4,0000	
	Median		4,0000	
	Variance		2,500	
	Std. Deviation		1,5811	
	Minimum		2,00	
	Maximum		6,00	
	Range		4,00	
	Interquartile Range		3,0000	
	Skewness		,000	,913
	Kurtosis		-1,200	2,000

10.5.2. Estimarea proporției p a unei populații

În SPSS nu este calculat direct intervalul de încredere pentru o proporție. Estimarea I.C. pentru o proporție folosind SPSS presupune efectuarea unui set de operații:

- Calculul estimației proporției
- Aflarea valorii variabilei Z
- Calcularea erorii standard
- Calculul limitelor I.C.

Se folosesc rezultatele afișate în fereastra output „Statistics“, obținut prin demersul: *Analyse* → *Descriptives Statistics* → *Frequencies*. Acestea sunt prezentate astfel:

Statistics VAR00001

N	Valid	8
	Missing	0
Std. Deviation		,4629
Variance		,2143

VAR00001

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	urban	6	75,0	75,0	75,0
	rural	2	25,0	25,0	100,0
	Total	8	100,0	100,0	



10.6. Alegeți răspunsul corect

1. Un estimator statistic reprezintă
 - a) o funcție de variabile aleatoare de selecție, independente și identic distribuite
 - b) o valoare tipică de sondaj
 - c) o estimație a parametrului populației
2. O estimație reprezintă
 - a) o funcție de variabile aleatoare de selecție, independente și identic distribuite
 - b) o valoare tipică de sondaj
 - c) o valoare a unui estimator al parametrului θ
3. Valoarea medie calculată la nivelul unui eșantion reprezintă
 - a) un parametru
 - b) o valoare tipică de sondaj
 - c) un estimator
 - d) o estimație
4. Valoarea medie calculată la nivelul unei populații reprezintă
 - a) un parametru
 - b) o valoare tipică de sondaj
 - c) un estimator
 - d) o estimație
5. Eroarea limită este folosită la stabilirea
 - a) preciziei unei estimații
 - b) măsurarea bias-ului
 - c) intervalului de variație a mediei de selecție față de media generală
6. Variabilele aleatoare de selecție, $X_1, X_2, \dots, X_n$,
 - a) sunt independente stochastice în ansamblu
 - b) urmează aceeași lege de probabilitate cu variabila studiată X
 - c) sunt dependente de probabilitatea de extragere a eșantioanelor
7. Un estimator statistic reprezintă
 - a) o funcție de n variabile aleatoare ($X_1, X_2, \dots, X_n$) independente și identic distribuite $\hat{\theta} = f(X_1, X_2, \dots, X_n)$

- b) o valoare a unei funcții $\hat{\theta}$ a parametrului θ
- c) o valoare tipică de sondaj de forma: $f(x_1, x_2, \dots, x_n)$

8. Un estimator este nedeplasat dacă:

- a) $M(\hat{\theta}) = \theta$
- b.) $V(\hat{\theta}) \rightarrow \theta$ când $n \rightarrow N$
- c) este convergent și are varianța cea mai mică posibil, față de varianța oricărui alt estimator calculat pentru același eșantion

9. Un estimator este convergent dacă:

- a) $M(\hat{\theta}) = \theta$
- b) $V(\hat{\theta}) \rightarrow \theta$ când $n \rightarrow N$
- c) are varianța cea mai mică posibil, față de varianța oricărui alt estimator calculat pentru același eșantion

10. Un estimator este eficient dacă:

- a) $M(\hat{\theta}) = \theta$
- b) $V(\hat{\theta}) \rightarrow \theta$ când $n \rightarrow N$
- c) este convergent și are varianța cea mai mică posibil față de varianța oricărui alt estimator calculat pentru același eșantion

11. Estimatorul $\hat{\mu}$ este

- a) nedeplasat, convergent, eficient
- b) deplasat, convergent, eficient
- c) nedeplasat, neconvergent, eficient

12. Estimatorul $\hat{p}$ este

- a) nedeplasat, convergent, eficient
- b) deplasat, convergent, eficient
- c) nedeplasat, neconvergent, eficient

13. Caracteristicile distribuției mediei de selecție, pentru o populație infinită, sunt

- a) $M(\hat{\mu}) = \mu$ și $V(\hat{\mu}) = \sigma^2$
- b) $M(\hat{\mu}) = \mu$ și $V(\hat{\mu}) = \sigma_{\mu}^2$
- c) $M(\hat{\mu}) = \mu$ și $V(\hat{\mu}) = \frac{\sigma^2}{n}$

14. Caracteristicile distribuției mediei de selecție, pentru o populație finită, sunt

- a) $M(\hat{\mu}) = \mu$ și $V(\hat{\mu}) = \sigma^2$
- b) $M(\hat{\mu}) = \mu$ și $V(\hat{\mu}) = \sigma_{\mu}^2$

$$c) M(\hat{\mu}) = \mu \text{ și } V(\hat{\mu}) = \frac{\sigma^2}{n} \cdot \frac{N-n}{N}$$

15. Caracteristicile distribuției proporției de selecție, pentru o populație infinită, sunt

$$a) M(\hat{p}) = p \text{ și } V(\hat{p}) = \sigma^2$$

$$b) M(\hat{p}) = p \text{ și } V(\hat{p}) = \sigma_p^2$$

$$c) M(\hat{p}) = p \text{ și } V(\hat{p}) = \frac{p(1-p)}{n}$$

16. Caracteristicile distribuției proporției de selecție, pentru o populație finită, sunt

$$a) M(\hat{p}) = p \text{ și } V(\hat{p}) = \sigma^2$$

$$b) M(\hat{p}) = p \text{ și } V(\hat{p}) = \sigma_p^2$$

$$c) M(\hat{p}) = p \text{ și } V(\hat{p}) = \frac{p(1-p)}{n} \cdot \frac{N-n}{N}$$

17. Mărimea intervalului de încredere al unui parametru crește o dată cu

a) reducerea volumului eșantionului

b) creșterea coeficientului de încredere

c) reducerea dispersiei

18. Mărimea intervalului de încredere al unui parametru crește o dată cu scăderea

a) volumului eșantionului

b) coeficientului de încredere

c) varianței populației

19. Mărimea intervalului de încredere al unei medii depinde de

a) volumul eșantionului

b) coeficientul de încredere

c) dispersia populației

20. Când nu se cunoaște varianța populației, limita inferioară a intervalului de încredere pentru estimarea mediei unei populații, în cazul eșantioanelor de volum $n < 30$, este

$$a) (\bar{x} - t_{\alpha} \cdot \frac{s'}{\sqrt{n}})$$

$$b) (\bar{x} - t_{\alpha/2} \cdot \frac{s'}{\sqrt{n}})$$

$$c) (\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}})$$

21. În cazul eşantioanelor de volum $n > 30$, când se cunoaște varianța populației, limita inferioară a intervalului de încredere al mediei unei populații este

a) $(\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}})$

b) $(\bar{x} - z_{\alpha} \cdot \frac{\sigma}{\sqrt{n}})$

c) $(\bar{x} - z_{\alpha/2} \cdot \frac{s}{\sqrt{n}})$



10.7. Recunoașteți exemplul corect

1. În 95 de cazuri dintr-o sută se poate afirma că media populației se găsește în intervalul
 - a. $(\bar{x} \pm 1,96 \cdot \sigma_{\bar{\mu}})$
 - b. $(\bar{x} \pm 1,28 \cdot \sigma_{\bar{\mu}})$
 - c. $(\bar{x} \pm 0,80 \cdot \sigma_{\bar{\mu}})$
2. În 80 de cazuri dintr-o sută se poate afirma că media populației se găsește în intervalul
 - a. $(\bar{x} \pm 1,94 \cdot \sigma_{\bar{\mu}})$
 - b. $(\bar{x} \pm 1,28 \cdot \sigma_{\bar{\mu}})$
 - c. $(\bar{x} \pm 0,80 \cdot \sigma_{\bar{\mu}})$
3. În 75 de cazuri dintr-o sută se poate afirma că media populației se găsește în intervalul
 - a. $(\bar{x} \pm 1,96 \cdot \sigma_{\bar{\mu}})$
 - b. $(\bar{x} \pm 1,28 \cdot \sigma_{\bar{\mu}})$
 - c. $(\bar{x} \pm 1,15 \cdot \sigma_{\bar{\mu}})$
4. Pentru un sondaj aleator simplu repetat, realizat asupra unui eșantion de volum $n=100$, valoarea mediei eșantionului și valoarea abaterii medii pătratice sunt: $\bar{x}=30 \text{ ani}$, $s=3 \text{ ani}$. Pentru un $\alpha=0,05$, valoarea erorii limită de selecție este

- a. 0,3
 - b. 0,588
 - c. 30

$$\Delta_{\bar{\mu}} = z \cdot \frac{s}{\sqrt{n}} = 1,96 \cdot \frac{3}{10} = 0,588$$
5. Pentru un sondaj aleator simplu repetat, realizat asupra unui eșantion de volum $n=9$, valoarea mediei eșantionului și valoarea abaterii medii pătratice sunt: $\bar{x}=30 \text{ ani}$, $s=3 \text{ ani}$. Pentru un $\alpha=0,05$, valoarea erorii limită de selecție este

- a. 2,262
 - b. 2,306
 - c. 2,228

$$\Delta_{\bar{\mu}} = t_{0,025, n-1} \cdot \frac{s}{\sqrt{n}} = 2,306 \cdot \frac{3}{3} = 2,306$$

6. Pentru un sondaj aleator simplu nerepetat, realizat asupra unui eșantion de volum $n=100$, extras dintr-o populație de volum $N=1000$, s-au obținut: $\bar{x} = 40$ ani, $s = 22$ ani. Pentru un $\alpha = 0,05$, valoarea erorii medii de selecție este

- a. 2,2
b. 2,0856
c. 4,312

$$\Delta_{\bar{x}} = \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}} = \frac{22}{10} \cdot \sqrt{\frac{900}{1000}} = 2,087$$

7. Pentru un sondaj aleator simplu nerepetat, realizat asupra unui eșantion de volum $n=25$, extras dintr-o populație de volum $N=1000$, s-au obținut: $\bar{x} = 40$ ani, $s^2 = 49$. Pentru un $\alpha = 0,10$, valoarea erorii limită de selecție este

- a. 2,364
b. 2,085
c. 1,4

$$\Delta_{\bar{x}} = t_{\alpha/2, n-1} \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}} = 2,364$$

8. Se extrage un eșantion aleatoriu simplu repetat de volum $n=900$ persoane. S-au obținut următoarele rezultate: $\bar{x} = 5$ mil.lei, $s = 3$ mil.lei. Considerând un risc de 5%, să se afle limita inferioară a intervalului de încredere pentru media populației din care s-a extras eșantionul.

- a. 4,732
b. 4,653
c. 4,804

$$z = 1,96$$

$$\bar{x} - z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} = 5 - 1,96 \cdot \frac{3}{\sqrt{900}} = 4,653$$

9. Se extrage un eșantion aleatoriu simplu repetat de volum $n=900$ persoane. S-au obținut următoarele rezultate: $\bar{x} = 5$ mil.lei, $s = 3$ mil.lei. Considerând un risc de 5%, să se afle limita superioară a intervalului de încredere pentru media populației din care s-a extras eșantionul.

- a. 5,365
b. 5,379
c. 5,196

$$\bar{x} + z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} = 5 + 1,96 \cdot \frac{3}{\sqrt{900}} = 5,3196$$

10. Se extrage un eșantion aleatoriu simplu repetat de volum $n=900$ persoane. S-au obținut următoarele rezultate: $\bar{x} = 5$ mil.lei, $s = 3$ mil.lei. Considerând un risc de 5%, să se afle limitele intervalului de încredere pentru media populației din care s-a extras eșantionul.

- a. (4,732; 5,653)
b. (4,653; 5,789)
c. (4,804; 5,196)

11. Se extrage un eșantion aleatoriu simplu repetat de volum $n = 10$ persoane. În urma înregistrării vârstei persoanei și a prelucrării datelor în SPSS s-au obținut următoarele rezultate:

Descriptives

			Statistic	Std. Error
VARSTA	Mean		25,9000	,5667
95% Confidence Interval for Mean	Lower Bound		24,6181	
	Upper Bound		27,1819	
	5% Trimmed Mean		26,0000	
	Median		26,0000	
	Variance		3,211	
	Std. Deviation		1,7920	
	Minimum		22,00	
	Maximum		28,00	
	Range		6,00	
	Interquartile Range		2,2500	
	Skewness		-,973	,687
	Kurtosis		1,484	1,334

Pe baza datelor din output, se poate considera că

- a. vârsta medie a populației din care a fost extras eșantionul este acoperită de intervalul (25,90; 24,6181), în condițiile unui risc de 5%
- b. vârsta medie a populației din care a fost extras eșantionul este acoperită de intervalul (24,6181; 27,1819), în condițiile unui risc de 95%
- c. vârsta medie a populației din care a fost extras eșantionul este acoperită de intervalul (24,6181; 27,1819), în condițiile unui risc de 5%

12. Se extrage un eșantion aleatoriu simplu repetat de volum $n=10$ persoane. În urma înregistrării vârstei persoanei și a prelucrării datelor în SPSS s-au obținut următoarele rezultate:

Descriptives

			Statistic	Std. Error
VARSTA	Mean		25,9000	,5667
95% Confidence Interval for Mean	Lower Bound		24,6181	
	Upper Bound		27,1819	
5% Trimmed Mean			26,0000	
Median			26,0000	
Variance			3,211	
Std. Deviation			1,7920	
Minimum			22,00	
Maximum			28,00	
Range			6,00	
Interquartile Range			2,2500	
Skewness			-,973	,687
Kurtosis			1,484	1,334

Pe baza datelor din output, se poate considera că

- a. vârsta medie calculată la nivelul eșantionului este 25,9 ani
 - b. vârsta medie calculată la nivelul eșantionului este 24,6181 ani
 - c. vârsta medie calculată la nivelul eșantionului este 27,1819 ani
13. Dintr-o populație grupată după sex s-au extras 2 eșantioane, $n_1=5$ bărbați, $n_2=5$ femei. În urma înregistrării vârstei persoanei și a prelucrării datelor în SPSS s-au obținut următoarele rezultate:

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
VARSTA Equal variances assumed	5,633	,045	6,200	8	,000	6,2000	1,0000	3,8940	8,5060
Equal variances not assumed			6,200	5,269	,001	6,2000	1,0000	3,6684	8,7316

- Pe baza datelor din output, se poate considera că
- a. diferența dintre vârstele medii ale celor două populații este acoperită de intervalul (3,6684 ani; 3,8940 ani), în condițiile unui risc de 5%
 - b. diferența dintre vârstele medii ale celor două populații este acoperită de intervalul (3,6684 ani; 8,7316 ani), în condițiile unui risc de 5%
 - c. diferența dintre vârstele medii ale celor două populații este acoperită de intervalul (3,6684 ani; 8,7316 ani), în condițiile unui risc de 95%

Statistics	VARSTA
Mean	25.9000
Std. Deviation	5.1819
Minimum	18.00
Maximum	32.00
Sum	129.50
N	5
Sum of Squares	130.50
Std. Error	1.0000
Lower Bound	3.8940
Upper Bound	8.5060
95% Confidence Interval	
Mean Difference	6.2000
Std. Error	1.0000
Lower Bound	3.6684
Upper Bound	8.7316
95% Confidence Interval	



10.8. Răspunsuri corecte

10.6. Alegeți răspunsul corect

1. a) Un estimator statistic reprezintă o funcție de variabile aleatoare de selecție independente și identic distribuite.
2. b) și c) O estimație reprezintă o valoare tipică de sondaj și o valoare a unui estimator al parametrului θ .
3. b) și d) Valoarea medie calculată la nivelul unui eșantion reprezintă o valoare tipică de sondaj, respectiv o estimație.
4. a) Valoarea medie calculată la nivelul unei populații reprezintă un parametru.
5. c) Eroarea limită este folosită la stabilirea intervalului de variație a mediei de selecție față de media generală.
6. a) și b) Variabilele aleatoare de selecție, $X_1, X_2, \dots, X_n$, sunt independente stochastic în ansamblu și urmează aceeași lege de probabilitate cu variabila studiată X .
7. a) Un estimator statistic reprezintă o funcție de n variabile aleatoare ($X_1, X_2, \dots, X_n$) independente și identic distribuite $\hat{\theta} = f(X_1, X_2, \dots, X_n)$.
8. a) Un estimator este nedeplasat dacă $M(\hat{\theta}) = \theta$.
9. b) Un estimator este convergent dacă $V(\hat{\theta}) \rightarrow 0$ când $n \rightarrow N$.
10. c) Un estimator este eficient dacă este convergent și are varianța cea mai mică posibil față de varianța oricărui alt estimator calculat pentru același eșantion.
11. a) Estimatorul $\hat{\mu}$ este nedeplasat, convergent, eficient.
12. a) Estimatorul $\hat{p}$ este nedeplasat, convergent, eficient.
13. b) și c) Caracteristicile distribuției mediei de selecție, pentru o populație infinită, sunt $M(\hat{\mu}) = \mu$ și $V(\hat{\mu}) = \sigma_{\mu}^2 = \frac{\sigma^2}{n}$.
14. b) și c) Caracteristicile distribuției mediei de selecție, pentru o populație finită, sunt: $M(\hat{\mu}) = \mu$ și $V(\hat{\mu}) = \sigma_{\mu}^2 = \frac{\sigma^2}{n} \cdot \frac{N-n}{N}$.

15. b) și c) Caracteristicile distribuției proporției de selecție, pentru o populație infinită, sunt: $M(\hat{p}) = p$ și $V(\hat{p}) = \sigma_{\hat{p}}^2 = \frac{p(1-p)}{n}$.

16. b) și c) Caracteristicile distribuției proporției de selecție, pentru o populație finită, sunt: $M(\hat{p}) = p$ și $V(\hat{p}) = \sigma_{\hat{p}}^2 = \frac{p(1-p)}{n} \cdot \frac{N-n}{N}$.

17. a) și b) Mărimea intervalului de încredere al unui parametru crește o dată cu reducerea volumului eșantionului, precum și cu creșterea coeficientului de încredere.

18. a) Mărimea intervalului de încredere al unui parametru crește o dată cu scăderea volumului eșantionului.

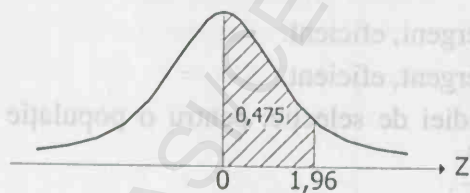
19. a), b) și c) Mărimea intervalului de încredere al unei medii depinde de volumul eșantionului, coeficientul de încredere și dispersia populației.

20. b) Când nu se cunoaște varianța populației, limita inferioară a intervalului de încredere pentru estimarea mediei unei populații, în cazul eșantioanelor de volum $n < 30$, este $(\bar{x} - t_{\alpha/2} \cdot \frac{s'}{\sqrt{n}})$.

21. a) În cazul eșantioanelor de volum $n > 30$, când se cunoaște varianța populației, limita inferioară a intervalului de încredere al mediei unei populații este $(\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}})$.

10.7. Recunoașteți exemplul corect

1. a) Valoarea funcției Laplace este $\phi(z) = \frac{0,95}{2} = 0,475$.



Pentru $\phi(z) = 0,475$, se citește în tabelul Gauss Laplace valoarea statisticii Z corespunzătoare, și anume: $z_i = 1,96$.

z	...	0.06	...
⋮			
1.9	...	0,475	...
⋮			

În 95 de cazuri dintr-o sută se poate afirma că media populației se găsește în intervalul $(\bar{x} \pm 1,96 \cdot \sigma_{\hat{\mu}})$.

2. b) Valoarea funcției Laplace este $\phi(z) = \frac{0,80}{2} = 0,40$.

Pentru $\phi(z) = 0,40$, se citește în tabelul Gauss Laplace valoarea statisticii Z corespunzătoare, și anume: $z_1 = 1,28$.

z	...	0.08	...
⋮			
1.2	...	0,400	...
⋮			

În 80 de cazuri dintr-o sută se poate afirma că media populației se găsește în intervalul $(\bar{x} \pm 1,28 \cdot \sigma_{\hat{\mu}})$.

3. c) Valoarea funcției Laplace este $\phi(z) = \frac{0,75}{2} = 0,375$.

Pentru $\phi(z) = 0,375$ se citește în tabelul Gauss Laplace valoarea statisticii Z corespunzătoare, și anume: $z_1 = 1,15$.

z	...	0.05	...
⋮			
1.1	...	0,375	...
⋮			

În 75 de cazuri dintr-o sută se poate afirma că media populației se găsește în intervalul $(\bar{x} \pm 1,15 \cdot \sigma_{\hat{\mu}})$

4. b) Pentru un sondaj aleator simplu repetat, realizat asupra unui eșantion de volum $n=100$, valoarea mediei eșantionului și valoarea abaterii medii pătratice sunt: $\bar{x} = 30 \text{ ani}$, $s = 3 \text{ ani}$. Pentru un $\alpha = 0,05$, valoarea erorii limită de

selecție este: $\Delta_{\mu} = z \cdot \sigma_{\mu} = z \cdot \frac{s}{\sqrt{n}} = 1,96 \cdot \frac{3}{\sqrt{100}} = 0,588$.

5. b) Pentru un sondaj aleator simplu repetat, realizat asupra unui eșantion de volum $n=9$, valoarea mediei eșantionului și valoarea abaterii medii pătratice sunt: $\bar{x} = 30 \text{ ani}$, $s = 3 \text{ ani}$. Pentru un $\alpha = 0,05$, valoarea erorii limită de

selecție este: $\Delta_{\mu} = t_{\alpha/2, n-1} \cdot \sigma_{\mu} = t_{0,025, 8} \cdot \frac{s}{\sqrt{n}} = 2,306 \cdot \frac{3}{\sqrt{9}} = 2,306$.

6. b) Pentru un sondaj aleator simplu nerepetat, realizat asupra unui eșantion de volum $n=100$, extras dintr-o populație de volum $N=1000$, s-au obținut: $\bar{x} = 40 \text{ ani}$, $s = 22 \text{ ani}$. Pentru un $\alpha = 0,05$, valoarea erorii medii de selecție

este: $\sigma_{\mu} = \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}} = \frac{22}{\sqrt{100}} \cdot \sqrt{\frac{1000-100}{1000}} = 2,0856$.

7. a) Pentru un sondaj aleator simplu nerepetat, realizat asupra unui eșantion de volum $n=25$, extras dintr-o populație de volum $N=1000$, s-au obținut:

$\bar{x} = 40 \text{ ani}$, $s^2 = 49$. Pentru un $\alpha = 0,10$, valoarea erorii limită de selecție este:

$\Delta_{\mu} = t_{\alpha/2, n-1} \cdot \sigma_{\mu} \cdot \sqrt{\frac{N-n}{N}} = t_{0,05, 24} \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}} = 1,711 \cdot \frac{7}{\sqrt{25}} \cdot \sqrt{\frac{1000-25}{1000}} = 2,364$.

8. c) Se extrage un eșantion aleatoriu simplu repetat de volum $n=900$ persoane. S-au obținut următoarele rezultate: $\bar{x} = 5 \text{ mil. lei}$, $s = 3 \text{ mil. lei}$. Considerând un risc de 5%, limita inferioară a intervalului de încredere pentru media populației

din care s-a extras eșantionul este: $\left(\bar{x} - z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \right) = \left(5 - 1,96 \cdot \frac{3}{\sqrt{900}} \right) = 4,804$.

9. c) Se extrage un eșantion aleatoriu simplu repetat de volum $n=900$ persoane. S-au obținut următoarele rezultate: $\bar{x} = 5 \text{ mil. lei}$, $s = 3 \text{ mil. lei}$. Considerând un risc de 5%, limita superioară a intervalului de încredere pentru media populației din care s-a extras eșantionul este:

$\left(\bar{x} + z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \right) = \left(5 + 1,96 \cdot \frac{3}{\sqrt{900}} \right) = 5,196$.

10. c) Se extrage un eșantion aleatoriu simplu repetat de volum $n=900$ persoane. S-au obținut următoarele rezultate: $\bar{x} = 5 \text{ mil. lei}$, $s = 3 \text{ mil. lei}$.

Considerând un risc de 5%, limitele intervalului de încredere pentru media populației din care s-a extras eșantionul sunt:

$$\left(\bar{x} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}\right) = \left(5 \pm 1,96 \cdot \frac{3}{\sqrt{900}}\right) = (4,804; 5,196).$$

11. c) Se extrage un eșantion aleatoriu simplu repetat de volum n=10 persoane. În urma înregistrării vârstei persoanei și a prelucrării datelor în SPSS s-au obținut următoarele rezultate:

Descriptives			Statistic	Std. Error
VARSTA	Mean		25,9000	,5667
	95% Confidence Interval for Mean	Lower Bound	24,6181	
		Upper Bound	27,1819	
	5% Trimmed Mean		26,0000	
	Median		26,0000	
	Variance		3,211	
	Std. Deviation		1,7920	
	Minimum		22,00	
	Maximum		28,00	
	Range		6,00	
	Interquartile Range		2,2500	
	Skewness		-,973	,687
	Kurtosis		1,484	1,334

Pe baza datelor din output, se poate considera că vârsta medie a populației din care a fost extras eșantionul este acoperită de intervalul (24,6181; 27,1819), în condițiile unui risc de 5%.

12. a) Se extrage un eșantion aleatoriu simplu repetat de volum n=10 persoane. În urma înregistrării vârstei persoanei și a prelucrării datelor în SPSS s-au obținut următoarele rezultate:



Interpretare: Probabilitatea ca o valoare a variabilei X să fie mai mică decât 250 este egală cu 0,977.

Descriptives

			Statistic	Std. Error
VARSTA	Mean		25,9000	,5667
	95% Confidence Interval for Mean	Lower Bound	24,6181	
		Upper Bound	27,1819	
	5% Trimmed Mean		26,0000	
	Median		26,0000	
	Variance		3,211	
	Std. Deviation		1,7920	
	Minimum		22,00	
	Maximum		28,00	
	Range		6,00	
	Interquartile Range		2,2500	
	Skewness		-,973	,687
	Kurtosis		1,484	1,334

Pe baza datelor din output, se poate considera că vârsta medie calculată la nivelul eșantionului este 25,9 ani.

13. b) Dintr-o populație grupată după sex s-au extras 2 eșantioane, $n_1=5$ bărbați, $n_2=5$ femei. În urma înregistrării vârstei persoanei și a prelucrării datelor în SPSS, s-au obținut următoarele rezultate:

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
VARSTA	Equal variances assumed	5,633	,045	6,200	8	,000	6,2000	1,0000	3,8940	8,5060
	Equal variances not assumed			6,200	5,269	,001	6,2000	1,0000	3,6684	8,7316

Pe baza datelor din output, se poate considera că diferența dintre vârstele medii ale celor două populații este acoperită de intervalul (3,6684 ani; 8,7316 ani), în condițiile unui risc de 5%.



10.9. Probleme

10.9.1. Calculul probabilității pentru o distribuție statistică

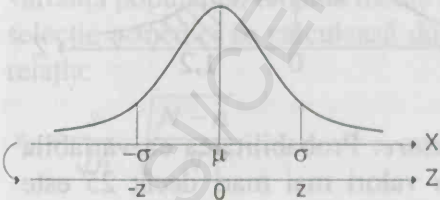
1. Pentru o variabilă X , care urmează o lege normală, definită de: $\mu = 200$ și $\sigma^2 = 625$, se cere să se afle probabilitatea $P(X < 250)$.

Pentru a afla probabilitatea unei variabile X este necesară, mai întâi, transformarea variabilei X într-o variabilă Z , apoi, citirea probabilității corespunzătoare pentru valoarea variabilei Z .

Adică:

1. Variabila $X \sim N(\mu, \sigma^2)$ se transformă în variabila $Z \sim N(0, 1)$, după relația:

$$z_i = \frac{x_i - \mu}{\sigma}.$$



2. Pentru o valoare calculată a variabilei Z , se citește din tabelul Gauss-Laplace probabilitatea corespunzătoare.

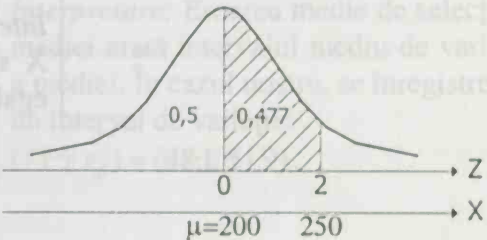
Valoarea variabilei centrate reduse Z , pentru $X=250$ este:

$$z_i = \frac{250 - 200}{25} = 2$$

Pentru $z_i=2$ se citește în tabelul Gauss Laplace valoarea funcției $\phi(z) = 0,477$.

z	0,00	...	...
$\vdots$			
2	0,477	...	...
$\vdots$			

Probabilitatea căutată este:
 $P(X < 250) = P(Z < 2) = 0,5 + 0,477 = 0,977$.



Interpretare: Probabilitatea ca o valoare a variabilei X să fie mai mică decât 250 este egală cu 0,977.

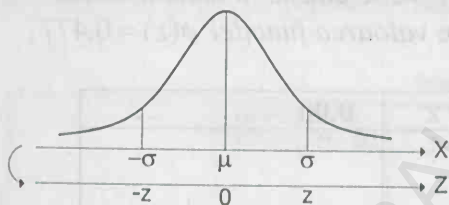
2. Pentru o variabilă X , care urmează o lege normală, definită de $\mu = 20$ și $\sigma^2 = 16$, se cere să se afle probabilitatea $P(X > 25)$.

Pentru a afla probabilitatea unei variabile X , este necesară, mai întâi, transformarea variabilei X într-o variabilă Z , apoi, citirea probabilității corespunzătoare pentru valoarea variabilei Z .

Adică:

1. Variabila $X \sim N(\mu, \sigma^2)$ se transformă în variabila $Z \sim N(0, 1)$, după relația:

$$z_i = \frac{x_i - \mu}{\sigma}.$$



2. Pentru o valoare calculată a variabilei Z , se citește din tabelul Gauss-Laplace probabilitatea corespunzătoare.

Valoarea variabilei centrate reduse Z pentru $X=25$ este:

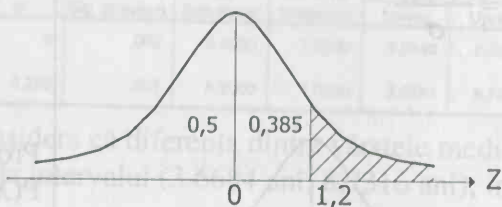
$$z_i = \frac{25 - 20}{4} = 1,2$$

Pentru valoarea $z_i=1,2$ se citește în tabelul Gauss Laplace valoarea funcției $\phi(z) = 0,385$.

z	0,00	...	...
⋮			
1.2	0,385	...	...
⋮			

Probabilitatea căutată este:

$$P(X > 25) = P(Z > 1,2) = 0,5 - 0,385 = 0,115.$$



Interpretare: Probabilitatea ca variabila X să ia valori mai mari decât 25 este egală cu 0,115.

10.9.2. Estimarea erorii medii de selecție a mediei

1. În urma observării și prelucrării datelor la nivelul unui eșantion extras aleator simplu repetat, de volum $n=100$, s-au obținut următoarele valori: $\bar{x} = 50$ milioane lei, $s=20$ milioane lei. Se cere să se estimeze eroarea medie de selecție a mediei.

În cazul unui sondaj aleator simplu repetat, când nu se cunoaște varianța populației, eroarea medie de selecție a mediei se calculează după relația:

$$s_{\bar{\mu}} = \frac{s}{\sqrt{n}}.$$

Eroarea medie de selecție a mediei:

$$s_{\bar{\mu}} = \frac{20}{\sqrt{100}} = 2 \text{ milioane lei}$$

Interpretare: Eroarea medie de selecție a mediei arată intervalul mediu de variație a mediei. În cazul nostru, se înregistrează un interval de variație: $(\bar{x} \pm s_{\bar{\mu}}) = (48; 52)$.

2. În urma observării și prelucrării datelor la nivelul unui eșantion de volum $n = 100$, extras aleator simplu nerepetat dintr-o populație formată din $N=1000$ persoane, s-au obținut următoarele valori: $\bar{x} = 50$ milioane lei, $s=20$ milioane lei. Se cere să se estimeze eroarea medie de selecție a mediei.

În cazul unui sondaj aleator simplu nerepetat, când nu se cunoaște varianța populației, eroarea medie de selecție a mediei se calculează după relația:

$$s_{\bar{\mu}} = \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}}.$$

Eroarea medie de selecție a mediei:

$$s_{\bar{\mu}} = \frac{20}{\sqrt{100}} \cdot \sqrt{\frac{1000-100}{1000}} = 1,9 \text{ milioane lei}$$

Interpretare: Eroarea medie de selecție a mediei arată intervalul mediu de variație a mediei. În cazul nostru, se înregistrează un interval de variație: $((\bar{x} \pm s_{\bar{\mu}}) = (48,1; 51,9))$.

10.9.3. Estimarea erorii limită a mediei

1. În urma observării și prelucrării datelor la nivelul unui eșantion extras aleator simplu repetat, de volum $n=100$, s-au obținut următoarele valori: $\bar{x}=55$ ani, $s=25$ ani. Se cere să se estimeze eroarea limită a mediei, considerând un risc de 0,05.

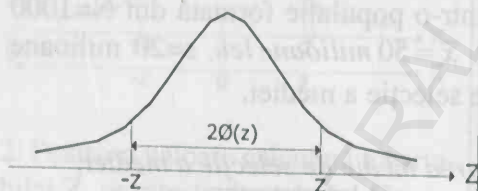
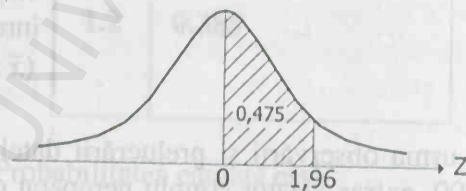
În cazul unui eșantion extras aleator simplu repetat, de volum $n>30$, când nu se cunoaște varianța populației, eroarea limită se calculează după relația:

$$\Delta_{\bar{\mu}} = z \cdot \sigma_{\bar{\mu}} = z \cdot \frac{s}{\sqrt{n}}.$$

Valoarea statisticii $Z \sim N(0,1)$ corespunzătoare unui risc ales, se citește din tabelul Gauss Laplace.

Pentru un risc $\alpha=0,05$, se citește din tabelul Gauss Laplace pentru

$$\phi(z) = \frac{0,95}{2} = 0,475, \text{ valoarea statisticii } Z, \text{ și anume : } z_i = 1,96.$$



z	... 0.06 ...
...	...
1.9	... 0,475 ...
...	...

Eroarea limită a mediei este:

$$\Delta_{\bar{\mu}} = 1,96 \cdot \frac{25}{\sqrt{100}} = 4,9 \text{ ani}$$

Interpretare: Pentru un risc asumat $\alpha=0,05$, ne putem aștepta ca eroarea limită (maximă admisibilă) a mediei să fie de $\pm 4,9$ ani.

2. În urma observării și prelucrării datelor la nivelul unui eșantion de volum $n=100$, extras aleator simplu nerepetat dintr-o populație de volum $N=1000$, s-au obținut următoarele valori: $\bar{x} = 55$ ani, $s=25$ ani. Se cere să se estimeze eroarea limită a mediei, considerând un risc de 0,05.

În cazul unui eșantion extras aleator simplu nerepetat, de volum $n > 30$, când nu se cunoaște varianța populației, eroarea limită se calculează după relația:

$$\Delta_{\bar{\mu}} = z \cdot \sigma_{\bar{\mu}} = z \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}}$$

Valoarea statisticii $Z \sim N(0,1)$ corespunzătoare unui risc ales, se citește din tabelul Gauss Laplace.

Pentru un risc $\alpha=0,05$, se citește din tabelul Gauss Laplace, pentru

$$\phi(z) = \frac{0,95}{2} = 0,475, \text{ valoarea statisticii } Z, \text{ și anume: } z_1 = 1,96.$$

Eroarea limită a mediei este:

$$\Delta_{\bar{\mu}} = 1,96 \cdot \frac{25}{\sqrt{100}} \sqrt{\frac{1000-100}{1000}} = 4,65 \text{ ani}$$

Interpretare: Pentru un risc asumat $\alpha=0,05$, ne putem aștepta ca eroarea limită (maximă admisibilă) a mediei să fie de $\pm 4,65$ ani.

3. La nivelul unui eșantion prelevat aleator simplu repetat, de volum $n=10$ persoane, s-a înregistrat distribuția după vârstă. În urma prelucrării datelor, s-au obținut: $\sum_i (x_i - \bar{x})^2 = 36$. Se cere să se estimeze eroarea limită a mediei, considerând un risc de 0,05.

În cazul unui eșantion extras aleator simplu repetat, de volum $n < 30$, când nu se cunoaște varianța populației, eroarea limită se calculează după relația:

$$\Delta_{\bar{\mu}} = t_{\alpha/2} \cdot \sigma_{\bar{\mu}} = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}},$$

unde:

$$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}}$$

Estimația abaterii standard este:

$$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{36}{10-1}} = 2 \text{ ani}$$

Pentru un risc de 0,05 și $v=n-1=9$ grade de libertate, se citește în tabelul Student valoarea statisticii

$$t_{0,025,9} = 2,262.$$

$t_{\alpha/2}$ este valoarea statisticii t care urmează o distribuție de probabilitate Student cu $(n-1)$ grade de libertate, $t \sim t(n-1)$.

α v	...	$t_{0.025}$	...
...	...	...	...
9	...	2,262	...
...	...	...	...

Eroarea limită a mediei de selecție este:

$$\Delta_{\mu} = 2,262 \cdot \frac{2}{\sqrt{100}} = 0,45$$

Interpretare: Pentru un risc asumat $\alpha=0,05$, ne putem aștepta ca eroarea limită (maximă admisibilă) a mediei să fie de $\pm 0,45$ ani.

10.9.4. Estimarea mediei prin interval de încredere

1. La nivelul unui eșantion extras aleator simplu repetat, de volum $n=5$ persoane, s-a înregistrat salariul lunar realizat (milioane lei): 5, 4, 6, 3, 2. Să se estimeze salariul mediu lunar pe o persoană, la nivelul ansamblului angajaților din care a fost extras eșantionul observat. Se garantează rezultatele cu o probabilitate de 0,95.

Când volumul eșantionului este $n < 30$ și nu se cunoaște varianța populației, atunci intervalul de încredere (I.C.) al mediei populației se calculează după relația:

$$\left(\bar{x} \pm t_{\alpha/2, n-1} \cdot \frac{s}{\sqrt{n}} \right),$$

unde:

$\bar{x}$ este media eșantionului:

Salariul mediu calculat la nivelul eșantionului este:

$$\bar{x} = \frac{5+4+6+3+2}{5} = 4 \text{ mil. lei}$$

Abaterea standard la nivelul eșantionului este:

$$s = \sqrt{\frac{(5-4)^2 + (4-4)^2 + \dots + (2-4)^2}{5-1}} = 1,41 \text{ mil. lei}$$

$$\bar{x} = \frac{\sum_i x_i}{n};$$

s este abaterea standard la nivelul eșantionului:

$$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n - 1}};$$

$t_{\alpha/2, n-1}$ este valoarea statisticii t care urmează o distribuție de probabilitate Student cu $(n-1)$ grade de libertate, $t \sim t(n-1)$.

Pentru un risc $\alpha = 0,05$ și $v = n-1 = 4$ grade de libertate se citește în tabelul Student valoarea statisticii $t_{0,025,4} = 2,776$.

α v	...	$t_{0,025}$	...
⋮			
4	...	2,776	...
⋮			

Intervalul de încredere este:

$$\left(4 \pm 2,776 \cdot \frac{1,41}{\sqrt{5}} \right) = (2,25 ; 5,75)$$

Interpretare: Cu o probabilitate de 0,95 se poate considera că salariul mediu al angajaților la nivelul populației din care s-a extras eșantionul este acoperit de intervalul: (2,25 milioane lei; 5,75 milioane lei). Ne asumăm un risc de 0,05 sau 5% ca valoarea adevărată a salariului mediu să nu fie acoperită de acest interval.

2. În urma observării și prelucrării datelor la nivelul unui eșantion extras aleator simplu repetat, de volum $n=300$, s-au obținut următoarele valori: $\bar{x} = 500\$$ și $s = 25\$$. Se cere să se estimeze prin interval de încredere media populației din care s-a extras eșantionul, considerând un risc de 0,05.

Când volumul eșantionului este $n > 30$ și nu se cunoaște varianța populației, atunci intervalul de încredere (I.C.) al mediei populației se calculează după relația:

Pentru un risc $\alpha = 0,05$, se citește în tabelul Gauss Laplace valoarea statisticii Z pentru $\phi(z) = \frac{0,95}{2} = 0,475$, și anume: $z_1 = 1,96$.

$$\left(\bar{x} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \right),$$

unde:

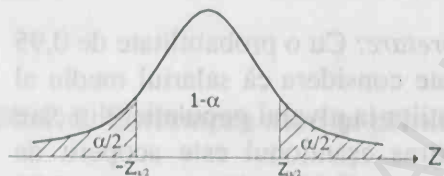
$\bar{x}$ este media eșantionului:

$$\bar{x} = \frac{\sum_i x_i}{n};$$

s este abaterea standard la nivelul eșantionului:

$$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}};$$

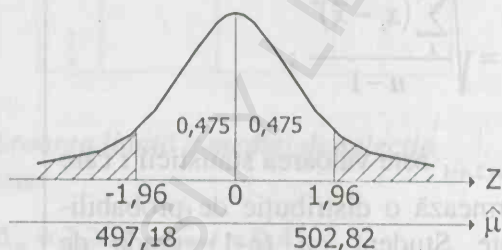
$z_{\alpha/2}$ este valoarea statisticii $Z \sim N(0,1)$ pentru un risc $\alpha/2$.



Valoarea statisticii Z , corespunzătoare unui risc ales, se citește din tabelul Gauss Laplace.

Intervalul de încredere al mediei populației este:

$$\left(500 \pm 1,96 \cdot \frac{25}{\sqrt{300}} \right) = (497,18; 502,82)$$



Interpretare: Cu o probabilitate de 0,95 se poate considera că media la nivelul populației este acoperită de intervalul (497,18\$; 502,82\$). Ne asumăm un risc de 0,05 sau 5% ca media populației să nu fie acoperită de acest interval.

3. Printr-un sondaj aleator simplu nerepetat se prelevă un eșantion de volum $n = 100$ dintr-o populație de volum $N=1000$ persoane. Vârsta medie la nivelul eșantionului observat este de $\bar{x} = 50$ ani, cu $s^2=25$. Estimați prin interval de încredere media la nivelul populației, considerând un risc $\alpha=0,05$.

Când volumul eșantionului este $n > 30$ și nu se cunoaște varianța populației, în cazul unui sondaj aleator simplu nerepetat, atunci

Pentru un risc $\alpha=0,05$, se citește în tabelul Gauss Laplace valoarea statisticii Z pentru $\phi(z) = \frac{0,95}{2} = 0,475$, și

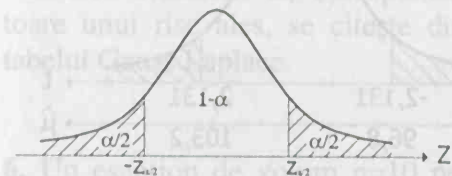
intervalul de încredere (I.C.) al mediei populației se calculează după relația:

$$\left(\bar{x} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}} \right),$$

unde:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}};$$

$z_{\alpha/2}$ este valoarea statisticii $Z \sim N(0,1)$ pentru un risc $\alpha/2$.



Valoarea statisticii Z, corespunzătoare unui risc ales, se citește din tabelul Gauss Laplace.

4. Cererea unui articol s-a înregistrat pe parcursul a 16 de zile. Prin prelucrarea datelor s-au obținut următoarele rezultate: $\bar{x} = 100$ produse, $s^2 = 36$. Să se estimeze prin interval de încredere cererea medie la nivelul unui an, considerându-se un risc de 0,05.

Când volumul eșantionului este $n < 30$ și nu se cunoaște varianța populației, atunci intervalul de încredere al mediei populației se calculează după relația:

$$\left(\bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \right),$$

unde:

anume: $z_i = 1,96$.

Intervalul de încredere este:

$$\left(50 \pm 1,96 \cdot \frac{\sqrt{25}}{\sqrt{100}} \cdot \sqrt{\frac{1000-100}{1000}} \right) = (49,12; 50,88)$$

Interpretare: Cu o probabilitate de 0,95 se poate considera că vârsta medie (μ) la nivelul populației este acoperită de intervalul: $49,12 \approx 49$ ani și $50,88 \approx 51$ ani.

Pentru un risc de 0,05 și $\nu = 16 - 1 = 15$ grade de libertate, se citește în tabelul Student valoarea statisticii

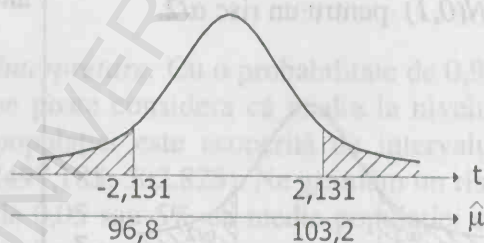
$$t_{0,025,15} = 2,131.$$

$t_{\alpha/2, n-1}$ este valoarea statisticii t care urmează o distribuție de probabilitate Student cu $(n-1)$ grade de libertate, $t \sim t(n-1)$.

$\alpha \backslash v$	...	$t_{0.025}$	...
...	...	...	...
15	...	2,131	...
...	...	...	...

Intervalul de încredere este:

$$\left(100 \pm 2,131 \cdot \frac{\sqrt{36}}{\sqrt{16}} \right) = (96,8 ; 103,2)$$



Interpretare: Cu o probabilitate de 0,95 se poate considera că cererea medie a produsului la nivelul anului este acoperită de intervalul: $96,8 \approx 97$ bucăți și $103,2 \approx 103$ bucăți.

5. Cererea unui articol s-a înregistrat pe parcursul a 36 de zile. Prin prelucrarea datelor eșantionului observat s-au obținut următoarele rezultate:

$$\bar{x} = 100 \text{ produse}, s^2 = 25.$$

Să se estimeze prin interval de încredere cererea medie la nivelul unui an, considerându-se un risc de 0,05.

Când volumul eșantionului este $n > 30$ și nu se cunoaște varianța populației, atunci intervalul de încredere al mediei populației se calculează după relația:

$$\left(\bar{x} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \right),$$

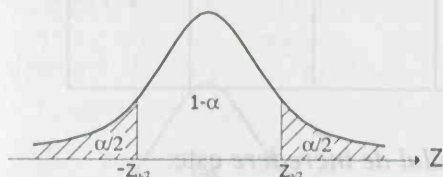
Pentru un risc de 0,05 se citește în tabelul Gauss Laplace valoarea statisticii $z_{\alpha/2} = 1,96$.

Intervalul de încredere este:

unde:

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}};$$

$z_{\alpha/2}$ este valoarea statisticii $Z \sim N(0,1)$ pentru un risc $\alpha/2$.



Valoarea statisticii Z , corespunzătoare unui risc ales, se citește din tabelul Gauss Laplace.

6. Un eșantion de volum $n=10$ persoane prelevat aleator simplu repetat, s-a înregistrat după vârstă și, prin prelucrarea datelor, s-au obținut următoarele rezultate: $\sum x_i = 250$; $\sum (x_i - \bar{x})^2 = 36$. Se cere să se estimeze prin interval de încredere vârsta medie a întregii populații din care a fost extras eșantionul, considerând un risc $\alpha=0,10$.

Când volumul eșantionului este $n < 30$ și nu se cunoaște varianța populației, atunci intervalul de încredere al mediei populației se calculează după relația:

$$\left(\bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \right),$$

unde:

$$\bar{x} = \frac{\sum x_i}{n};$$

$$\left(100 \pm 1,96 \cdot \frac{5}{\sqrt{36}} \right) = (98,37; 101,63)$$

Interpretare: Cu o probabilitate de 0,95 se poate considera că cererea medie a produsului la nivelul anului este acoperită de intervalul: $98,37 \approx 98$ bucăți și $101,63 \approx 102$ bucăți.

Vârsta medie calculată la nivelul eșantionului observat este:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{250}{10} = 25 \text{ ani}$$

Abaterea standard la nivelul eșantionului este:

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{36}{10-1}} = 2 \text{ ani}$$

Pentru un risc de 0,10 și $v=n-1=9$ grade de libertate se citește în tabelul

$$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}};$$

$t_{\alpha/2, n-1}$ este valoarea statisticii t care urmează o distribuție de probabilitate Student cu $(n-1)$ grade de libertate, $t \sim t(n-1)$.

Student valoarea statisticii

$$t_{0.05, 9} = 1,833.$$

α	...	$t_{0.05}$	...
v			
...			
9	...	1,833	...
...			

Intervalul de încredere este:

$$\left(25 \pm 1,833 \cdot \frac{2}{\sqrt{10}} \right) = (23,84; 26,16)$$

Interpretare: Cu o probabilitate de 0,90 se poate considera că vârsta medie a populației din care a fost extras eșantionul observat este acoperită de intervalul: $(23,84 \approx 24 \text{ ani}; 26,16 \approx 26 \text{ ani})$.

7. Pentru un eșantion de 101 studenți s-a înregistrat timpul consumat necesar introducerii datelor pe calculator pentru realizarea unei lucrări. În urma prelucrării datelor, s-au obținut următoarele rezultate privind realizarea lucrării:

$$\bar{x} = 70 \text{ minute}, \quad \sum_{i=1}^{101} (x_i - \bar{x})^2 = 1400.$$

Să se estimeze prin interval de încredere timpul mediu consumat de un student din populația din care s-a extras eșantionul, garantând rezultatele cu o probabilitate de 0,98.

Când volumul eșantionului este $n > 30$ și nu se cunoaște varianța populației, atunci intervalul de încredere al mediei populației se calculează după relația:

$$\left(\bar{x} \pm z_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \right),$$

unde:

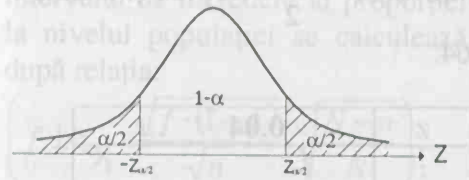
Abaterea standard la nivelul eșantionului este:

$$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{1400}{101-1}} = 3,74$$

Pentru un risc de 0,02, se citește în tabelul Gauss Laplace valoarea

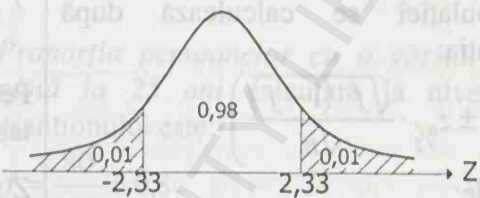
$$\bar{x} = \frac{\sum_i x_i}{n};$$
$$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}};$$

$z_{\alpha/2}$ este valoarea statisticii $Z \sim N(0,1)$ pentru un risc $\alpha/2$.



Valoarea statisticii Z, corespunzătoare unui risc ales, se citește din tabelul Gauss Laplace.

statisticii Z pentru $\phi(z) = \frac{0,98}{2} = 0,49$,
și anume : $z_i = 2,31$.



z	...	0.01	...
⋮			
2.3	...	0,49	...
⋮			

I.C. pentru media populației este:

$$\left(70 \pm 2,31 \cdot \frac{3,74}{\sqrt{101}} \right) = (69,15 ; 70,85)$$

Interpretare: Cu o probabilitate de 0,98 se poate considera că timpul mediu consumat de un student din populația din care s-a extras eșantionul pentru elaborarea unei lucrări este acoperit de intervalul: (69,15 minute; 70,85 minute).

10.9.5. Estimarea proporției prin interval de încredere

1. Pentru un eșantion format din 100 persoane, extras aleator simplu și repetat, se constată că numărul persoanelor cu o vârstă de până la 25 ani este de 80. Se cere să se estimeze prin interval de încredere proporția acestei categorii de vârstă, la nivelul întregii populații din care a fost extras eșantionul, considerând un risc $\alpha = 0,10$.

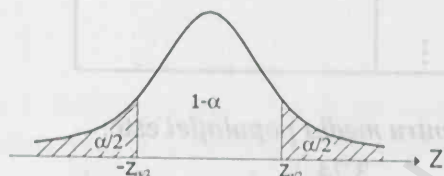
Când volumul eșantionului este $n > 30$ și nu se cunoaște varianța populației, atunci intervalul de încredere al proporției la nivelul populației se calculează după relația:

$$\left(f \pm z_{\alpha/2} \cdot \frac{\sqrt{f \cdot (1-f)}}{\sqrt{n}} \right),$$

unde:

$$f = \frac{n_A}{n},$$

$z_{\alpha/2}$ este valoarea statisticii $Z \sim N(0,1)$ pentru un risc $\alpha/2$.



Valoarea statisticii Z , corespunzătoare unui risc ales, se citește din tabelul Gauss Laplace.

Proporția persoanelor cu o vârstă de până la 25 ani, calculată la nivelul eșantionului, este:

$$f = \frac{80}{100} = 0,8$$

Pentru un risc de 0,10, se citește în tabelul Gauss Laplace valoarea statisticii

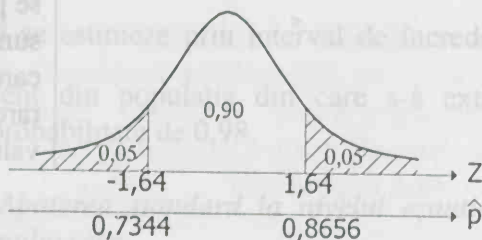
$$Z \text{ pentru } \phi(z) = \frac{0,90}{2} = 0,45, \text{ și anume:}$$

$$z_i = 1,64.$$

z	...	0.04	...
...			
1.6	...	0,449	...
...			

Intervalul de încredere este:

$$\left(0,8 \pm 1,64 \cdot \frac{\sqrt{0,8 \cdot (1-0,8)}}{\sqrt{100}} \right) = (0,7344; 0,8656)$$



Interpretare: Cu o probabilitate de 0,90 se poate considera că proporția persoanelor cu o vârstă de până la 25 ani la nivelul populației este acoperită de intervalul (73,44% ; 86,56%).

2. Pentru un eșantion format din 100 persoane, extras aleator simplu și nerepetat dintr-o populație de volum $N=1000$ unități, se constată că numărul persoanelor cu o vârstă de până la 25 ani este de 80. Se cere să se estimeze prin interval de încredere proporția acestei categorii de vârstă la nivelul întregii populații din care a fost extras eșantionul, considerând un risc $\alpha=0,10$.

Când volumul eșantionului este $n>30$ și nu se cunoaște varianța populației, în cazul unui sondaj aleator simplu nerepetat, atunci intervalul de încredere al proporției la nivelul populației se calculează după relația:

$$\left(f \pm z_{\alpha/2} \cdot \frac{\sqrt{f \cdot (1-f)}}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}} \right),$$

unde:

$$f = \frac{n_A}{n},$$

$z_{\alpha/2}$ este valoarea statisticii

$Z \sim N(0,1)$ pentru un risc $\alpha/2$. Valoarea statisticii Z , corespunzătoare unui risc ales, se citește din tabelul Gauss Laplace.

Proporția persoanelor cu o vârstă de până la 25 ani calculată la nivelul eșantionului este:

$$f = \frac{80}{100} = 0,8.$$

Pentru un risc de 0,10, se citește în tabelul Gauss Laplace *valoarea statisticii*

Z pentru $\phi(z) = \frac{0,90}{2} = 0,45$, și anume :

$$z_i = 1,64.$$

Intervalul de încredere pentru proporție este:

$$\left(0,8 \pm 1,64 \cdot \frac{\sqrt{0,8 \cdot (1-0,8)}}{\sqrt{100}} \cdot \sqrt{\frac{1000-100}{1000}} \right) = (0,738; 0,862)$$

Interpretare: Cu o probabilitate de 0,90 se poate considera că proporția persoanelor cu o vârstă de până la 25 ani la nivelul populației este acoperită de intervalul: (73,8% ; 86,2%).

3. Dintr-o populație de volum $N=1000$, se prelevă aleator simplu repetat un eșantion de volum $n=100$. Ponderea categoriei A observată în eșantion este $f=30\%$. Să se estimeze prin interval de încredere proporția acestei categorii la nivelul populației, considerând un risc $\alpha=0,05$.

Când volumul eșantionului este $n>30$ și nu se cunoaște varianța populației, atunci intervalul de încredere al proporției la nivelul populației se calculează după relația:

Pentru un risc $\alpha=0,05$, se citește în tabelul Gauss Laplace *valoarea statisticii* Z pentru

$$\phi(z) = \frac{0,95}{2} = 0,475, \text{ și anume: } z_i = 1,96.$$

$$\left(f \pm z_{\alpha/2} \cdot \frac{\sqrt{f \cdot (100 - f)}}{\sqrt{n}} \right),$$

unde:

$$f = \frac{n_A}{n} \cdot 100,$$

$z_{\alpha/2}$ este valoarea statisticii

$Z \sim N(0,1)$ pentru un risc $\alpha/2$. Valoarea statisticii Z , corespunzătoare unui risc ales, se citește din tabelul Gauss Laplace.

I.C. pentru proporție este:

$$\left(30 \pm 1,96 \cdot \frac{\sqrt{30 \cdot (100 - 30)}}{\sqrt{100}} \right) = (21,02; 38,98)$$

Interpretare: Cu o probabilitate de 0,95 se poate considera că proporția persoanelor cu o vârstă de până la 25 ani la nivelul populației este acoperită de intervalul (21,02% ; 38,98%).

4. Dintr-o populație de volum $N=1000$, se prelevă aleator simplu repetat un eșantion de volum $n=20$. Ponderea categoriei A observată în eșantion este $f=40\%$. Să se estimeze prin interval de încredere proporția acestei categorii la nivelul populației, considerând un risc $\alpha=0,05$.

Când volumul eșantionului este $n < 30$ și nu se cunoaște varianța populației, atunci intervalul de încredere al proporției la nivelul populației se calculează după relația:

$$\left(f \pm t_{\alpha/2} \cdot \frac{\sqrt{f \cdot (100 - f)}}{\sqrt{n}} \right),$$

unde:

$$f = \frac{n_A}{n} \cdot 100,$$

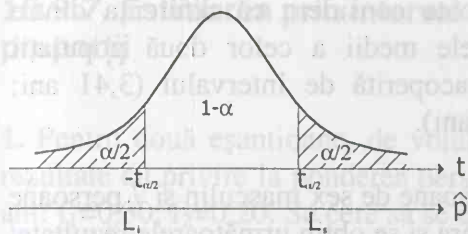
$t_{\alpha/2}$ este valoarea statisticii t care urmează o distribuție de probabilitate Student cu $(n-1)$ grade de libertate, $t \sim t(n-1)$.

Pentru un risc $\alpha=0,05$, se citește în tabelul Student *valoarea statisticii t* corespunzătoare, $t_{0,025,19}=2,093$.

α v	...	$t_{0,025}$	...
...			
19	...	2,093	...
...			

I.C. pentru proporție este:

$$\left(40 \pm 2,093 \cdot \frac{\sqrt{40 \cdot (100 - 40)}}{\sqrt{20}} \right) = (17,08; 62,92)$$



Valoarea statisticii t , corespunzătoare unui risc ales, se citește din tabelul Student.

Interpretare: Cu o probabilitate de 0,95 se poate considera că proporția persoanelor cu o vârstă de până la 25 ani la nivelul populației este acoperită de intervalul (17,08%; 62,92%).

10.9.6. Estimarea prin interval de încredere a diferenței dintre două medii

1. Se extrag aleator simplu repetat două eșantioane formate fiecare din 100 de persoane. Se înregistrează vârsta persoanelor și se obțin următoarele rezultate: $\bar{x}_1 = 42$ ani, $\bar{x}_2 = 37$ ani; $s_1 = 4$ ani, $s_2 = 7$ ani. Să se calculeze intervalul de încredere pentru diferența dintre vârstele medii ale celor două populații din care au fost extrase eșantioanele observate. Se cere garantarea rezultatului cu o probabilitate de 0,95.

Când eșantioanele observate sunt de volum mare ($n_1 > 30$ și $n_2 > 30$) și nu se cunoaște varianța populației, atunci intervalul de încredere pentru diferența dintre două medii se calculează după relația:

$$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

unde:

$\bar{x}_1, \bar{x}_2$ reprezintă mediile la nivelul eșantioanelor;

s_1^2, s_2^2 reprezintă abaterile medii pătratice la nivelul eșantioanelor,

$z_{\alpha/2}$ este valoarea statisticii

$Z \sim N(0,1)$ pentru un risc $\alpha/2$.

Pentru un risc $\alpha=0,05$, se citește în tabelul Gauss Laplace valoarea statisticii Z pentru

$$\phi(z) = \frac{0,95}{2} = 0,475, \text{ și anume: } z_i = 1,96.$$

În exemplul considerat, $n_1 = n_2 = 100$.

I.C. pentru diferența mediilor populațiilor este:

$$(42 - 37) \pm 1,96 \cdot \sqrt{\frac{16}{100} + \frac{49}{100}} = (3,41; 6,59)$$

Interpretare: Cu o probabilitate de 0,95

Valoarea statisticii Z, corespunzătoare unui risc ales, se citește din tabelul Gauss Laplace.

se poate considera că diferența dintre vârstele medii a celor două populații este acoperită de intervalul (3,41 ani; 6,59 ani).

2. Pentru două eșantioane formate din 5 persoane de sex masculin și 7 persoane de sex feminin se înregistrează vârsta acestora și se obțin următoarele rezultate: $\bar{x}_M = 30\text{ ani}$, $\bar{x}_F = 26\text{ ani}$; $s_M^2 = 9$, $s_F^2 = 4$. Să se calculeze intervalul de încredere pentru diferența dintre vârstele medii ale populațiilor, pe cele două categorii. Se cere garantarea rezultatului cu o probabilitate de 0,95.

Când eșantioanele observate sunt de volum mic și nu se cunoaște varianța populației, atunci intervalul de încredere pentru diferența dintre două medii se calculează după relația:

$$(\bar{x}_M - \bar{x}_F) \pm t_{\alpha/2} \cdot \sqrt{\frac{s_M^2}{n_1} + \frac{s_F^2}{n_2}},$$

unde:
 $\bar{x}_M, \bar{x}_F$ reprezintă mediile la nivelul eșantioanelor;
 s_M^2, s_F^2 reprezintă abaterile medii pătrățice la nivelul eșantioanelor;
 $t_{\alpha/2, n_1+n_2-2}$ este valoarea statisticii t Student pentru un risc $\alpha/2$ și (n_1+n_2-2) grade de libertate.

Pentru un risc $\alpha=0,05$ și $v=n_1+n_2-2=5+7-2=10$ grade de libertate, valoarea statisticii t Student este $t_{0,025,10}=2,228$.

α v	...	$t_{0.025}$	...
10	...	2,228	...

I.C. pentru diferența dintre mediile celor două populații este:
 $(30 - 26) \pm 2,228 \cdot \sqrt{\frac{9}{5} + \frac{4}{7}} =$
 $= (0,57; 7,43)$

Interpretare: Cu o probabilitate de 0,95 se poate considera că diferența dintre vârstele medii a celor două populații este acoperită de intervalul: (0,57 ani; 7,43 ani).

10.9.7. Estimarea prin interval de încredere a diferenței dintre două proporții

1. Pentru două eșantioane, de volum $n_1=100$, $n_2=200$, se cunosc următoarele rezultate cu privire la ponderea personalului angajat în vârstă de până la 25 de ani: $f_1=0,30$; $f_2=0,20$. Se cere să se estimeze prin interval de încredere diferența dintre proporții la nivelul populațiilor din care s-au extras eșantioanele, considerând un risc $\alpha=0,10$.

Când eșantioanele observate sunt de volum mare, atunci intervalul de încredere pentru diferența dintre două proporții se calculează după relația:

$$(f_1 - f_2) \pm z_{\alpha/2} \cdot \sqrt{\frac{f_1(1-f_1)}{n_1} + \frac{f_2(1-f_2)}{n_2}},$$

unde:

$z_{\alpha/2}$ este valoarea statisticii $Z \sim N(0,1)$ pentru un risc $\alpha/2$. Valoarea statisticii Z , corespunzătoare unui risc ales, se citește din tabelul Gauss Laplace.

Pentru un risc $\alpha=0,10$, se citește în tabelul Gauss Laplace valoarea statisticii Z pentru

$$\phi(z) = \frac{0,90}{2} = 0,45, \text{ și anume: } z_i=1,64.$$

I.C. pentru diferența dintre proporții la nivelul populațiilor este:

$$\begin{aligned} & (0,30 - 0,20) \pm \\ & \pm 1,64 \cdot \sqrt{\frac{0,30(1-0,30)}{100} + \frac{0,20(1-0,20)}{200}} = \\ & = (0,012; 0,188) \end{aligned}$$

Interpretare: Cu o probabilitate de 0,90 se poate considera că diferența dintre proporția personalului angajat în vârstă de până la 25 de ani a celor două populații este acoperită de intervalul (1,2% ; 18,8%).

2. Pentru două eșantioane, de volum $n_1=15$, $n_2=10$, se cunosc următoarele rezultate cu privire la ponderea personalului angajat în vârstă de până la 25 de ani: $f_1=0,60$; $f_2=0,20$. Se cere să se estimeze prin interval de încredere diferența dintre proporții la nivelul populațiilor din care s-au extras eșantioanele, considerând un risc $\alpha=0,05$.

Când eșantioanele observate sunt de volum mic, atunci intervalul de

Pentru un risc $\alpha=0,05$ și $v=n_1+n_2-2=15+10-2=23$ grade de libertate,

încredere pentru diferența dintre două proporții se calculează după relația:

$$(f_1 - f_2) \pm t_{\alpha/2} \cdot \sqrt{\frac{f_1(1-f_1)}{n_1} + \frac{f_2(1-f_2)}{n_2}}$$

unde:
 $t_{\alpha/2, n_1+n_2-2}$ este valoarea statisticii t Student pentru un risc $\alpha/2$ și (n_1+n_2-2) grade de libertate.

valoarea statisticii t Student este $t_{0,025,23}=2,069$.

α v	...	$t_{0.025}$	...
...	...	...	...
23	...	2,069	...
...	...	...	...

I.C. pentru diferența dintre proporții la nivelul populațiilor este:

$$\begin{aligned} &(0,60-0,20) \pm \\ &\pm 2,069 \cdot \sqrt{\frac{0,60(1-0,60)}{15} + \frac{0,20(1-0,20)}{10}} = \\ &= (0,03; 0,77) \end{aligned}$$

Interpretare: Cu o probabilitate de 0,95 se poate considera că diferența dintre ponderea personalului angajat în vârstă de până la 25 de ani a celor două populații este acoperită de intervalul: (3% ; 77%).

Capitolul 11

TESTAREA IPOTEZELOR STATISTICE

- ◆ Demersul testării unei ipoteze statistice
- ◆ Testarea ipotezelor asupra unui parametru
- ◆ Verificarea normalității unei distribuții
- ◆ Testarea ipotezelor statistice cu SPSS
- ◆ Alegeți răspunsul corect
- ◆ Recunoașteți exemplul corect
- ◆ Răspunsuri corecte
- ◆ Probleme

Testarea ipotezelor statistice este un procedeu prin care, în funcție de anumite reguli de decizie, se poate respinge sau nu o ipoteză admisă asupra unui parametru sau asupra unei distribuții.

11.1. Demersul testării unei ipoteze statistice

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 331-336)

Demersul testării unei ipoteze presupune parcurgerea unor etape și rezolvarea problemelor pe care le implică, și anume:

1. Se formulează ipotezele, în funcție de problema pusă;
2. Se alege un test statistic în funcție de distribuția de selecție a statisticii considerate. Se alege un estimator $\hat{\theta}$ al unui parametru θ de testat;
3. Se alege un prag de semnificație α pentru test;
4. Se stabilesc regulile de decizie, definind regiunile de „acceptare” și de „respingere” a ipotezei H_0 ;
5. Se calculează valoarea statisticii test, folosind datele înregistrate prin sondaj;
6. Se compară valoarea calculată a statisticii test cu valoarea teoretică;
7. Se ia decizia de a nu respinge sau de a respinge ipoteza admisă.

- **Ipoteze statistice**
 - Ipoteză nulă
 - Ipoteză alternativă

- **Erori de testare**
 - Eroarea de tip I
 - Eroarea de tip II

- **Teste asupra ipotezelor statistice**
 - Regiunea de respingere
 - Regiune de acceptare

- Testul bilateral
- Testul unilateral la dreapta
- Testul unilateral la stânga

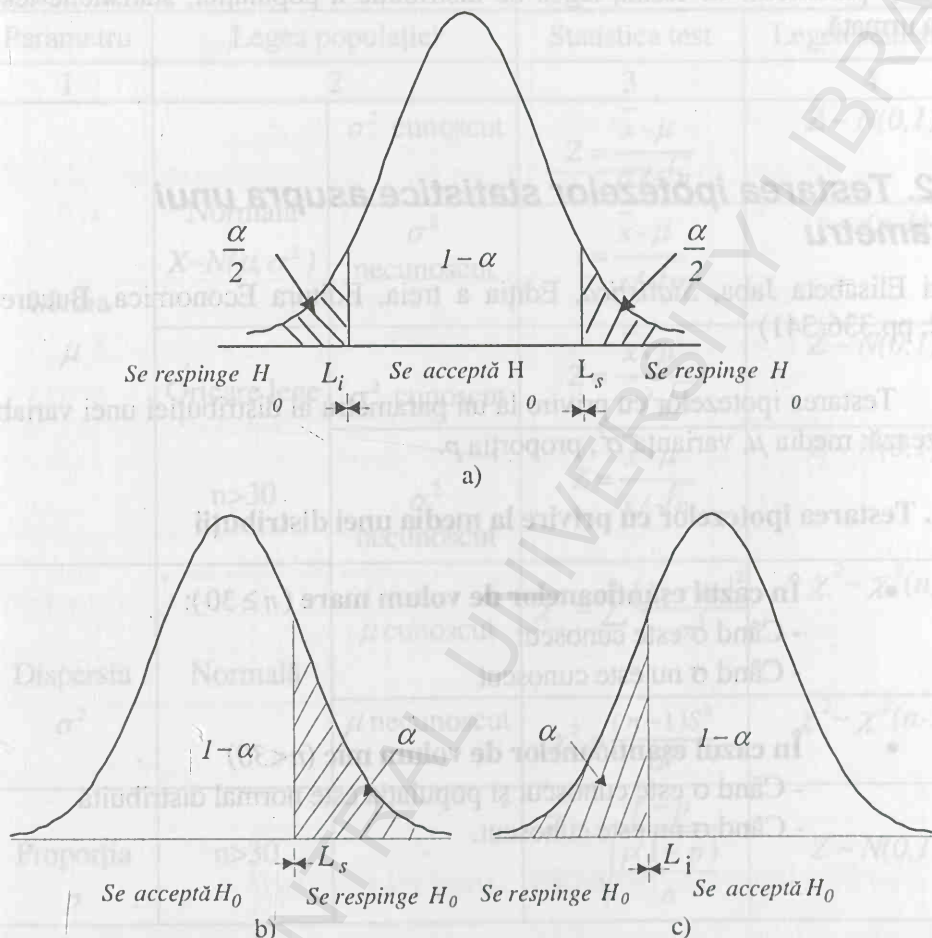


Fig. 11.1 Regiunea de acceptare și regiunea de respingere a ipotezei H_0

• Tipuri de teste

1. Teste parametrice

- testul t – testul Student
- testul χ^2
- testul F – testul Fisher

2. Teste neparametrice

- testul Wilcoxon (1945);
- testul Mann-Whitney (1947);
- testul Kolmogorov-Smirnov (1933).

În tabelul 11.1 sunt prezentate testele curent folosite în statistică. Sunt precizați parametrii de testat, legea de distribuție a populației, statisticile test și legea urmată.

11.2. Testarea ipotezelor statistice asupra unui parametru

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp.336-341)

Testarea ipotezelor cu privire la un parametru al distribuției unei variabile X vizează: media μ , varianța σ^2 , proporția p .

3.2.1. Testarea ipotezelor cu privire la media unei distribuții

- **În cazul eșantioanelor de volum mare ($n \geq 30$):**
 - Când σ este cunoscut
 - Când σ nu este cunoscut
- **În cazul eșantioanelor de volum mic ($n < 30$):**
 - Când σ este cunoscut și populația este normal distribuită
 - Când σ nu este cunoscut.

Tabelul 11.1 Statistici test folosite în testarea ipotezelor cu privire la un parametru

Parametru	Legea populației	Statistica test	Legea statisticii
1	2	3	4
Media μ	Normală $X \sim N(\mu, \sigma^2)$	σ^2 cunoscut $Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$	$Z \sim N(0,1)$
		σ^2 necunoscut $t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$	$t \sim t(n-1)$
	Oricare lege $n > 30$	σ^2 cunoscut $Z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$	$Z \sim N(0,1)$
		σ^2 necunoscut $Z = \frac{\bar{x} - \mu}{s / \sqrt{n}}$	$Z \sim N(0,1)$
Dispersia σ^2	Normală	μ cunoscut $\chi^2 = \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^2}$	$\chi^2 \sim \chi^2(n)$
		μ necunoscut $\chi^2 = \frac{(n-1)S^2}{\sigma^2}$	$\chi^2 \sim \chi^2(n-1)$
Proporția p	$n > 30$	- $Z = \frac{f - p}{\sqrt{\frac{p(1-p)}{n}}}$	$Z \sim N(0,1)$

3.2.2. Testarea ipotezelor privind compararea a două medii

În mod analog cu testarea ipotezelor privind un parametru se procedează în testarea ipotezelor în cazul comparării parametrilor a două distribuții. O sinteză a situațiilor de comparare a doi parametri și a statisticilor test adecvate este prezentată în tabelul 11.2.

Tabelul 11.2 Statistici test folosite în compararea a doi parametri

Parametru	Legea populației		Statistica test	Legea statisticii
1	2		3	4
Medii μ_1 și μ_2	Normale Oricare n_1 și n_2 >30	σ_1^2 și σ_2^2 cunoscute	$Z_{\mu_1-\mu_2} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$Z_{\mu_1-\mu_2} \sim N(0,1)$
	Normale n_1 și n_2 >30 Oricare n_1 și n_2 >30	σ_1^2 și σ_2^2 necunoscute	$Z_{\mu_1-\mu_2} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$	$Z_{\mu_1-\mu_2} \sim N(0,1)$
Proporții p_1 și p_2	n_1 și n_2 >30	-	$Z_{p_1-p_2} = \frac{f_1 - f_2}{\sqrt{\frac{f_1(1-f_1)}{n_1} + \frac{f_2(1-f_2)}{n_2}}}$	$Z_{p_1-p_2} \sim N(0,1)$

11.3. Verificarea normalității unei distribuții

(Vezi Elisabeta Jaba, Statistica, Ediția a treia, Editura Economica, București, 2002, pp. 341-351)

- Curba frecvențelor, histograma
- Testul Kolmogorov-Smirnov-Lilliefors
- Testul χ^2
- Testul Jarque-Bera
- Dreapta Henry

11.4. Testarea ipotezelor statistice cu SPSS

(Vezi Elisabeta Jaba și Ana Grama, *Analiza statistică cu SPSS sub Windows*, Editura Polirom, Iași, 2004, pp. 208-229)

11.4.1. Testarea egalității unei medii cu o valoare specificată

- **Demersul în SPSS**

Analyse → *Compare Means* → *One-Sample T Test*

- **Rezultatele în SPSS**

Rezultatele afișate în fereastra „*T-Test*” se prezintă astfel:

T-Test

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
VAR00001	10	6,4000	2,0111	,6360

One-Sample Test

	Test Value = 7				95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference		
					Lower	Upper
VAR00001	-,943	9	,370	-,6000	-2,0386	,8386

11.4.2. Testarea egalității mediilor a două eșantioane independente

- **Demersul în SPSS**

Analyse → *Compare Means* → *Independent-Sample T Test*

- **Rezultatele în SPSS**

Rezultatele afișate în output-ul „*T-Test*” se prezintă astfel:

T-Test
Group Statistics

	VAR00002	N	Mean	Std. Deviation	Std. Error Mean
VAR00001	masculin	5	4,8000	,4472	,2000
	feminin	5	8,0000	1,5811	,7071

Independent Samples Test

	Levene's Test for quality of Variance		t-test for Equality of Means							
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference		
								Lower	Upper	
VAR0001	Equal varian assumed	5,016	,055	-4,355	8	,002	-3,2000	,7348	-4,8946	-1,5054
	Equal varian not assumed			-4,355	4,636	,009	-3,2000	,7348	-5,1345	-1,2655

11.4.3. Testarea egalității mediilor a două eșantioane perechi

- Demersul în SPSS
Analyse → *Compare Means* → *Paired Sample T Test*
- Rezultatele în SPSS
Rezultatele afișate în output-ul „T-Test” se prezintă astfel:

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	VAR00001	6,4000	10	2,0111	,6360
	VAR00003	6,8000	10	2,0976	,6633

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	VAR00001 & VAR00003	10	,943	,000

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 VAR00001 - VAR00002	-,4000	,6992	,2211	-,9002	,1002	-1,809	9	,104

11.4.4. Testarea egalității unei proporții cu o valoare specificată

- Demersul în SPSS

Analyse → Compare Means → Binomial Test

- Rezultatele în SPSS

Rezultatele afișate în output-ul „NPar Tests” se prezintă astfel:

NPar Tests

Binomial Test

	Category	N	Observed Prop.	Test Prop.	Exact Sig. (2-tailed)
VAR00002	Group 1 masculin	3	,30	,50	,344
	Group 2 feminin	7	,70		
	Total	10	1,00		

11.4.5. Testarea egalității a două și mai multe proporții

- Demersul în SPSS

Analyse → Compare Means → Chi-Square Test

- Rezultatele în SPSS

Rezultatele afișate în output-ul „Chi-Square Test” se prezintă astfel:

Chi-Square Test

Frequencies

	VAR00002			
	Category	Observed N	Expected N	Residual
1	masculin	3	5,0	-2,0
2	feminin	7	5,0	2,0
Total		10		

Test Statistics

	VAR00002
Chi-Square ^a	,000
df	1
Asymp. Sig.	,206

- a. 0 cells (,0%) have expected frequencies less than 5. The minimum expected cell frequency is 5,0.



11.5. Alegeți răspunsul corect

1. Testul t se poate folosi pentru
 - a. testarea ipotezelor statistice cu privire la parametrii unui model
 - b. testarea normalității unei distribuții
 - c. testarea egalității a două medii
2. Eroarea de tip I este eroarea comisă în testarea ipotezelor
 - a. de a respinge ipoteza nulă când în realitate aceasta este adevărată
 - b. de a accepta ipoteza nulă atunci când aceasta este falsă
 - c. de a respinge ipoteza alternativă când în realitate aceasta este adevărată
3. Eroarea de tip II este eroarea comisă în testarea ipotezelor
 - a. de a respinge ipoteza nulă când în realitate aceasta este adevărată
 - b. de a accepta ipoteza nulă atunci când aceasta este falsă
 - c. de a accepta ipoteza alternativă când în realitate aceasta este falsă
4. Regiunea de respingere a unei ipoteze este intervalul dintr-o distribuție de probabilitate a unei statistici considerate în care
 - a. se respinge ipoteza nulă
 - b. se acceptă ipoteza nulă
 - c. se acceptă ipoteza alternativă
5. Regiunea de acceptare a unei ipoteze este intervalul dintr-o distribuție de probabilitate a unei statistici considerate în care
 - a. se respinge ipoteza nulă
 - b. se acceptă ipoteza nulă
 - c. se acceptă ipoteza alternativă
6. Într-un test bilateral, în legătură cu un parametru θ , ipotezele care pot fi formulate sunt:
 - a. $H_0: \theta = \theta_0; H_1: \theta \neq \theta_0$
 - b. $H_0: \theta = \theta_0; H_1: \theta > \theta_0$
 - c. $H_0: \theta = \theta_0; H_1: \theta < \theta_0$
7. Într-un test unilateral la dreapta, în legătură cu un parametru θ , ipotezele care pot fi formulate sunt:
 - a. $H_0: \theta = \theta_0; H_1: \theta \neq \theta_0$

b. $H_o : \theta = \theta_o ; H_1 : \theta > \theta_o$

c. $H_o : \theta = \theta_o ; H_1 : \theta < \theta_o$

8. Într-un test unilateral la stânga, în legătură cu un parametru θ , ipotezele care pot fi formulate sunt:

a. $H_o : \theta = \theta_o ; H_1 : \theta \neq \theta_o$

b. $H_o : \theta = \theta_o ; H_1 : \theta > \theta_o$

c. $H_o : \theta = \theta_o ; H_1 : \theta < \theta_o$

9. Valoarea critică inferioară (L_i) este definită de relația:

a. $L_i = \mu_{\hat{\theta}} - z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$

b. $L_i = \mu_{\hat{\theta}} + z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$

c. $L_i = \mu_{\hat{\theta}} \pm z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$

10. Valoarea critică superioară (L_s) este definită de relația:

a. $L_s = \mu_{\hat{\theta}} - z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$

b. $L_s = \mu_{\hat{\theta}} + z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$

c. $L_s = \mu_{\hat{\theta}} \pm z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$

11. În testarea ipotezelor cu privire la media unei distribuții, în cazul eșanțioanelor de volum mare, când σ este cunoscut, statistica test se calculează după relația:

a. $z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

b. $z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$

c. $t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$

12. În testarea ipotezelor cu privire la media unei distribuții, în cazul eșanțioanelor de volum mare, când σ nu este cunoscut, statistica test se calculează după relația:

a. $z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

b. $z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$

c. $t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$

13. În testarea ipotezelor cu privire la media unei distribuții, în cazul eșantioanelor de volum mic, când σ nu este cunoscut, statistica test se calculează după relația:

a. $z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

b. $z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$

c. $t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$

14. În testarea ipotezelor cu privire la proporția unei distribuții, în cazul eșantioanelor de volum mare, statistica test se calculează după relația:

a. $z = \frac{f - p_0}{\sqrt{f(1-f)} / \sqrt{n}}$

b. $t = \frac{f - p_0}{\sqrt{f(1-f)} / \sqrt{n}}$

c. $t = \frac{f - p_0}{\sqrt{f(1-f)} / \sqrt{N}}$

15. În testarea ipotezelor cu privire la proporția unei distribuții, în cazul eșantioanelor de volum mic, statistica test se calculează după relația:

a. $z = \frac{f - p_0}{\sqrt{f(1-f)} / \sqrt{n}}$

b. $t = \frac{f - p_0}{\sqrt{f(1-f)} / \sqrt{n}}$

c. $t = \frac{f - p_0}{\sqrt{f(1-f)} / \sqrt{N}}$

16. Condiția de normalitate este cerută în aplicarea

a. testelor neparametrice

b. testelor parametrice

c. analizei de regresie

17. Condiția de homoscedasticitate este cerută în aplicarea

a. testelor neparametrice

b. ANOVA

c. analizei de regresie

18. Testarea proporției unei distribuții se poate realiza folosind:

a. testul χ^2

b. testul t Student

c. testul Z

19. Testarea mediei unei distribuții se poate realiza folosind:

a. testul χ^2

b. testul t Student

c. histograma

20. Pentru testarea egalității a două medii se folosește

a. testul t Student

b. testul χ^2

c. testul Fisher

21. Pragul de semnificație este

a. pragul până la care se admite un rezultat aproximativ

b. o limită a erorii de speța I

c. o limită a erorii de speța a II-a

22. În cazul folosirii unui test unilateral, riscul asumat de a respinge pe nedrept ipoteza H_0 este

a. $\alpha/2$

b. α

c. $(1-\alpha)$

23. În cazul testării unei ipoteze se pot produce erori de

a. acceptare pe nedrept a unei ipoteze false

b. respingere pe nedrept a unei ipoteze adevărate

c. acoperire a populației eșantionate

24. Verificarea normalității unei distribuții se poate realiza pe cale grafică folosind:

a. testul χ^2

b. dreapta Henry

c. histograma



11.6. Recunoașteți exemplul corect

1. În vederea testării semnificației diferenței dintre o medie μ a unei populații față de o valoare fixă μ_0 , pentru un eșantion de volum $n=50$, se obține o valoare calculată a statisticii test $z=2,45$. Valoarea teoretică este $z=1,64$. În cazul unui test bilateral, probabilitatea de a se respinge pe nedrept ipoteza de nul este:
a. 0,10
b. 0,05
c. 0,02
2. În vederea testării semnificației diferenței dintre o medie μ și o valoare fixă μ_0 , pentru un eșantion de volum $n=50$, se obține o valoare calculată a statisticii test $z=1,45$. În cazul unui test bilateral, dacă valoarea teoretică $z=1,96$, riscul asumat de a respinge pe nedrept ipoteza H_0 este:
a. $\alpha=0,01$
b. $\alpha=0,50$
c. $\alpha=0,95$
3. În vederea testării semnificației diferenței dintre o medie μ și o valoare fixă μ_0 , pentru un eșantion de volum $n=15$, se obține o valoare calculată a statisticii test $t=2,75$. Valoarea teoretică a statisticii t Student este $t_{0,01/14}=2,624$. Probabilitatea de a respinge pe nedrept ipoteza de nul, în cazul unui test unilateral, este
a. 0.05
b. 0.01
c. 0.02
4. În urma testării diferenței dintre două proporții, calculate la nivelul a două eșantioane independente, s-a obținut pentru statistica test t o valoare egală cu $t = 5,32$. Dacă se consideră un risc $\alpha=0,05$ și numărul gradelor de libertate $v=15$, atunci :
a. diferența dintre cele două proporții este semnificativă
b. se respinge ipoteza H_0
c. diferența dintre cele două proporții nu este semnificativă
5. În vederea testării ipotezei de egalitate a două medii, se extrag două eșantioane de volum $n_1=10$ și $n_2=15$. Valoarea teoretică a statisticii test, în acest caz, pentru un prag de semnificație de $\alpha=0.05$, se alege pentru:
a. $v=24$ grade de libertate

b. $v=23$ grade de libertate

c. $v=25$ grade de libertate

6. Pentru un eșantion de volum $n=150$, în vederea testării semnificației diferenței dintre o proporție p pentru o populație față de o valoare fixă p_0 , se obține o valoare calculată a statisticii test $z=2,42$. Garantând rezultatele cu o probabilitate de 90%, se poate afirma că:

a. se respinge ipoteza de egalitate dintre proporția populației p și valoarea fixă p_0 ;

b. se acceptă ipoteza de egalitate dintre proporția populației p și valoarea fixă p_0 ;

7. Folosind un eșantion de volum $n=15$ pentru testarea semnificației diferenței dintre o medie μ față de o valoare fixă μ_0 , valoarea calculată a statisticii test este $t=2,75$. Valoarea teoretică a statisticii t Student este $t_{0,01,14}=2,624$. Riscul asumat de a respinge pe nedrept ipoteza H_0 de egalitate între μ și μ_0 pentru un test bilateral, este

a. $\alpha=0,05$

b. $\alpha=0,01$

c. $\alpha=0,02$

8. În urma testării concordanței dintre o distribuție observată și o distribuție teoretică a rezultat o valoare $\chi^2 = 14,4$. Decizia corespunzătoare, considerând un risc $\alpha=0,05$ și numărul gradelor de libertate $v=6$, este:

a. distribuția observată urmează o lege normală

b. distribuția observată nu urmează o lege normală

c. distribuția observată urmează o lege Fisher

9. În urma testării concordanței dintre o distribuție observată și o distribuție teoretică, a rezultat un $\chi^2 = 14,4$. Decizia corespunzătoare, considerând un risc $\alpha=0,05$ și numărul gradelor de libertate $v=15$, este:

a. distribuția observată urmează o lege normală

b. distribuția observată nu urmează o lege normală

c. distribuția observată urmează o lege Fisher

10. Dintr-o populație grupată după sex s-au extras 2 eșantioane, $n_1=5$ bărbați, $n_2=5$ femei. În urma înregistrării vârstei persoanei și a prelucrării datelor în SPSS s-au obținut următoarele rezultate:

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference		Lower	Upper
VARSTA	Equal variances assumed	5,633	,045	6,200	8	,000	6,2000	1,0000		3,8940	8,5060
	Equal variances not assumed			6,200	5,269	,001	6,2000	1,0000		3,6684	8,7316

Pe baza datelor din output, se poate considera că

- se respinge ipoteza de egalitate a varianțelor populațiilor din care au fost extrase eșantioanele, în condițiile unui risc de 5%
- se respinge ipoteza de egalitate a varianțelor populațiilor din care au fost extrase eșantioanele, în condițiile unui risc de 95%
- se acceptă ipoteza de egalitate a varianțelor populațiilor din care au fost extrase eșantioanele, în condițiile unui risc de 5%

11. Dintr-o populație grupată după sex s-au extras 2 eșantioane, $n_1=5$ bărbați, $n_2=5$ femei. În urma înregistrării vârstei persoanei și a prelucrării datelor în SPSS s-au obținut următoarele rezultate:

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference		Lower	Upper
VARSTA	Equal variances assumed	5,633	,045	6,200	8	,000	6,2000	1,0000		3,8940	8,5060
	Equal variances not assumed			6,200	5,269	,001	6,2000	1,0000		3,6684	8,7316

Pe baza datelor din output, se poate considera că

- există diferențe semnificative între vârsta medie a persoanelor de sex masculin și vârsta medie a persoanelor de sex feminin, la nivelul populațiilor din care au fost extrase eșantioanele, în condițiile unui risc de 5%
- există diferențe semnificative între vârsta medie a persoanelor de sex masculin și vârsta medie a persoanelor de sex feminin, la nivelul populațiilor din care au fost extrase eșantioanele, în condițiile unui risc de 95%
- nu există diferențe semnificative între vârsta medie a persoanelor de sex masculin și vârsta medie a persoanelor de sex feminin, la nivelul populațiilor din care au fost extrase eșantioanele, în condițiile unui risc de 5%

12. În vederea testării ipotezei de egalitate a salariului mediu cu o valoare fixă, μ_0 , s-a înregistrat un eșantion de volum $n=10$ persoane. În urma prelucrării datelor în SPSS s-au obținut următoarele rezultate :

One-Sample Test

	Test Value = 46					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
VARSTA	-2,637	9	,027	-3,5000	-6,5020	-,4980

Pe baza datelor din output, se poate considera că

a. există diferențe semnificative între salariul mediu al populației din care a fost extras eșantionul observat și valoarea μ_0 , în condițiile unui risc de 5%

b. există diferențe semnificative între salariul mediu al populației din care a fost extras eșantionul observat și valoarea μ_0 , în condițiile unui risc de 95%

c. nu există diferențe semnificative între salariul mediu al populației din care a fost extras eșantionul observat și valoarea μ_0 , în condițiile unui risc de 5%

13. În vederea testării ipotezei de egalitate a salariului mediu cu o valoare fixă, μ_0 , s-a înregistrat un eșantion de volum $n=10$ persoane. În urma prelucrării datelor în SPSS s-au obținut următoarele rezultate :

One-Sample Test

	Test Value = 46					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
VARSTA	-2,637	9	,027	-3,5000	-6,5020	-,4980

În acest caz, se pot formula următoarele ipoteze statistice:

a. $H_0 : \mu \neq 46$; $H_1 : \mu = 46$

b. $H_0 : \mu = 46$; $H_1 : \mu < 46$

c. $H_0 : \mu = 46$; $H_1 : \mu \neq 46$

14. În vederea testării ipotezei de egalitate a salariului mediu cu o valoare fixă, μ_0 , s-a înregistrat un eșantion de volum $n=10$ persoane. În urma prelucrării datelor în SPSS s-au obținut următoarele rezultate:

One-Sample Test

	Test Value = 22					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
VARSTA	4,019	9	,003	1,7000	,7432	2,6568

Pe baza datelor din output, se poate considera că

- există diferențe semnificative între salariul mediu al populației din care a fost extras eșantionul observat și valoarea μ_0 , în condițiile unui risc de 5%
- există diferențe semnificative între salariul mediu al populației din care a fost extras eșantionul observat și valoarea μ_0 , în condițiile unui risc de 95%
- nu există diferențe semnificative între salariul mediu al populației din care a fost extras eșantionul observat și valoarea μ_0 , în condițiile unui risc de 5%

15. Pentru testarea ipotezei de normalitate a unei distribuții s-a aplicat testul χ^2 asupra datelor înregistrate pe un eșantion de volum $n=40$ persoane. În urma prelucrării datelor în SPSS s-au obținut următoarele rezultate :

Test Statistics

	VAR00001
Chi-Square ^a	13,700
df	5
Asymp. Sig.	,018

- 0 cells (,0%) have expected frequencies less than 5.
- The minimum expected cell frequency is 6,7.

Pe baza datelor din output, se poate considera că

- se respinge ipoteza de normalitate a distribuției, în condițiile unui risc de 5%
- se respinge ipoteza de normalitate a distribuției, în condițiile unui risc de 95%
- se acceptă ipoteza de normalitate a distribuției, în condițiile unui risc de 5%

16. Pentru testarea ipotezei de normalitate a unei distribuții s-a aplicat testul Kolmogorov-Smirnov asupra datelor înregistrate pe un eșantion de volum $n=40$

persoane. În urma prelucrării datelor în SPSS s-au obținut următoarele rezultate :

One-Sample Kolmogorov-Smirnov Test

		VAR00001
N		40
Normal Parameters ^{a,b}	Mean	6,2500
	Std. Deviation	2,4677
Most Extreme Differences	Absolute	,160
	Positive	,131
	Negative	-,160
Kolmogorov-Smirnov Z		1,010
Asymp. Sig. (2-tailed)		,260

a. Test distribution is Normal.

b. Calculated from data.

Pe baza datelor din output, se poate considera că

- se respinge ipoteza de normalitate a distribuției, în condițiile unui risc de 5%
- se respinge ipoteza de normalitate a distribuției, în condițiile unui risc de 95%
- se acceptă ipoteza de normalitate a distribuției, în condițiile unui risc de 5%



11.7. Răspunsuri corecte

11.5. Alegeți răspunsul corect

1. a) și c) Testul t se poate folosi pentru testarea ipotezelor statistice cu privire la parametrul unui model și pentru testarea egalității a două medii.
2. a) Eroarea de tip I este eroarea comisă în testarea ipotezelor de a respinge ipoteza nulă când în realitate aceasta este adevărată.
3. b) Eroarea de tip II este eroarea comisă în testarea ipotezelor de a accepta ipoteza nulă, atunci când aceasta este falsă.
4. a) și c) Regiunea de respingere a unei ipoteze este intervalul dintr-o distribuție de probabilitate a unei statistici considerate în care se respinge ipoteza nulă și se acceptă ipoteza alternativă.
5. b) Regiunea de acceptare a unei ipoteze este intervalul dintr-o distribuție de probabilitate a unei statistici considerate în care se acceptă ipoteza nulă.
6. a) Într-un test bilateral, în legătură cu un parametru θ , ipotezele care pot fi formulate sunt: $H_0: \theta = \theta_0$; $H_1: \theta \neq \theta_0$.
7. b) Într-un test unilateral la dreapta, în legătură cu un parametru θ , ipotezele care pot fi formulate sunt: $H_0: \theta = \theta_0$; $H_1: \theta > \theta_0$.
8. c) Într-un test unilateral la stânga, în legătură cu un parametru θ , ipotezele care pot fi formulate sunt: $H_0: \theta = \theta_0$; $H_1: \theta < \theta_0$.
9. a) Valoarea critică inferioară (L_i) este definită de relația: $L_i = \mu_{\hat{\theta}} - z_{\alpha/2} \cdot \sigma_{\hat{\theta}}$.
10. b) Valoarea critică superioară (L_s) este definită de relația:

$$L_s = \mu_{\hat{\theta}} + z_{\alpha/2} \cdot \sigma_{\hat{\theta}}.$$
11. a) În cazul eșantioanelor de volum mare, când σ este cunoscut, în testarea ipotezelor cu privire la media unei distribuții, statistica test se calculează după relația: $z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$.
12. b) În cazul eșantioanelor de volum mare, când σ nu este cunoscut, în testarea ipotezelor cu privire la media unei distribuții, statistica test se calculează după relația: $z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$.

13. c) În cazul eșantioanelor de volum mic, când σ nu este cunoscut, în testarea ipotezelor cu privire la media unei distribuții, statistica test se calculează după

$$\text{relația: } t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}.$$

14. a) În cazul eșantioanelor de volum mare, în testarea ipotezelor cu privire la proporția unei distribuții, statistica test se calculează după relația:

$$z = \frac{f - p_0}{\sqrt{f(1-f)/n}}$$

15. b) În cazul eșantioanelor de volum mic, în testarea ipotezelor cu privire la proporția unei distribuții, statistica test se calculează după relația:

$$t = \frac{f - p_0}{\sqrt{f(1-f)/n}}$$

16. b) și c) Condiția de normalitate este cerută în aplicarea testelor parametrice și a analizei de regresie.

17. b) Condiția de homoscedasticitate este cerută în aplicarea ANOVA.

18. b) și c) Testarea proporției unei distribuții se poate realiza folosind testul t Student și testul Z.

19. b) Testarea mediei unei distribuții se poate realiza folosind testul t Student.

20. a) Pentru testarea egalității a două medii se folosește testul t Student.

21. b) Pragul de semnificație este o limită a erorii de speța I.

22. b) În cazul folosirii unui test unilateral, riscul asumat de a respinge pe nedrept ipoteza H_0 este α .

23. a) și b) În cazul testării unei ipoteze se pot produce erori de acceptare pe nedrept a unei ipoteze false și de respingere pe nedrept a unei ipoteze adevărate.

24. b) și c) Verificarea normalității unei distribuții se poate realiza pe cale grafică folosind dreapta Henry și histograma.

11.6. Recunoașteți exemplul corect

1. a) În vederea testării semnificației diferenței dintre o medie μ a unei populații față de o valoare fixă μ_0 , pentru un eșantion de volum $n=50$, se obține o valoare calculată a statisticii test $z=2,45$. Valoarea teoretică este $z=1,64$. Probabilitatea de a se respinge pe nedrept ipoteza de nul, în cazul unui test bilateral, este 0,10.

2. b) În vederea testării semnificației diferenței dintre o medie μ și o valoare fixă μ_0 , pentru un eșantion de volum $n=50$, se obține o valoare calculată a

statisticii test $z=1,45$. În cazul unui test bilateral, dacă valoarea teoretică $z=1,96$, riscul asumat de a respinge pe nedrept ipoteza H_0 este $\alpha=0,05$.

3. b) În vederea testării semnificației diferenței dintre media unei populații μ și o valoare fixă μ_0 , pentru un eșantion de volum $n=15$, valoarea calculată a statisticii test este $t=2,75$. Valoarea teoretică a statisticii t Student este $t_{0,01,14}=2,624$. Probabilitatea de a respinge pe nedrept ipoteza de nul, în cazul unui test unilateral, este 0,01.

4. a) și b) În urma testării diferenței dintre două proporții calculate la nivelul a două eșantioane independente s-a obținut pentru statistica test t o valoare egală cu $t = -5,32$. Dacă se consideră un risc $\alpha=0,05$ și numărul gradelor de libertate $v=15$, atunci se respinge ipoteza H_0 , diferența dintre cele două proporții este semnificativă ($t_{0,025,15}=2,131$).

5. b) În vederea testării ipotezei de egalitate a două medii, se extrag două eșantioane de volum $n_1=10$ și $n_2=15$. Valoarea teoretică a statisticii test, în acest caz, pentru un prag de semnificație de $\alpha=0,05$, se alege pentru: $v = n_1 + n_2 - 2 = 23$ grade de libertate.

6. a) Pentru un eșantion de volum $n=150$, în vederea testării semnificației diferenței dintre o proporție p pentru o populație față de o valoare fixă p_0 , se obține o valoare calculată a statisticii test $z = 2,42$. Garantând rezultatele cu o probabilitate de 90%, se poate afirma că se respinge ipoteza de egalitate dintre proporția populației p și valoarea fixă p_0 ($z_{calc}=2,42 > z_{teor}=1,64$).

7. c) Folosind un eșantion de volum $n=15$ pentru testarea semnificației diferenței dintre o medie μ față de o valoare fixă μ_0 , valoarea calculată a statisticii test este $t=2,75$. Valoarea teoretică a statisticii t Student este $t_{0,01,14}=2,624$. Riscul asumat de a respinge pe nedrept ipoteza H_0 de egalitate între μ și μ_0 pentru un test bilateral, este $\alpha=0,02$.

8. b) În urma testării concordanței dintre o distribuție observată și o distribuție teoretică a rezultat o valoare $\chi^2 = 14,4$. Decizia corespunzătoare, considerând un risc $\alpha=0,05$ și numărul gradelor de libertate $v=6$, este: $(\chi^2 = 14,4) > (\chi^2_{0,05,6} = 12,592)$, se respinge ipoteza H_0 de normalitate.

Distribuția considerată nu urmează o lege normală.

9. a) În urma testării concordanței dintre o distribuție observată și o distribuție teoretică, a rezultat un $\chi^2 = 14,4$. Decizia corespunzătoare, considerând un risc $\alpha=0,05$ și numărul gradelor de libertate $v=15$, este: $(\chi^2 = 14,4) < (\chi^2_{0,05,15} = 24,996)$ se acceptă ipoteza H_0 de normalitate.

Distribuția considerată urmează o lege normală.

10. a) Dintr-o populație grupată după sex s-au extras 2 eșantioane, $n_1=5$ bărbați, $n_2=5$ femei. În urma înregistrării vârstei persoanei și a prelucrării datelor în SPSS s-au obținut următoarele rezultate:

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
VARSTA Equal variances assumed	5,633	,045	6,200	8	,000	6,2000	1,0000	3,8940	8,5060
Equal variances not assumed			6,200	5,269	,001	6,2000	1,0000	3,6684	8,7316

Valoarea Sig.=0,045<0,05, se respinge ipoteza de egalitate a varianțelor populațiilor din care au fost extrase eșantioanele, în condițiile unui risc de 5%.

11. a) Dintr-o populație grupată după sex, s-au extras 2 eșantioane, $n_1=5$ bărbați, $n_2=5$ femei. În urma înregistrării vârstei persoanei și a prelucrării datelor în SPSS s-au obținut următoarele rezultate:

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
VARSTA Equal variances assumed	5,633	,045	6,200	8	,000	6,2000	1,0000	3,8940	8,5060
Equal variances not assumed			6,200	5,269	,001	6,2000	1,0000	3,6684	8,7316

Valoarea Sig.=0.001<0,05, se respinge ipoteza de egalitate între mediile comparate. Pe baza datelor din output, se poate considera că există diferențe semnificative între vârsta medie a persoanelor de sex masculin și vârsta medie a persoanelor de sex feminin, la nivelul populațiilor din care au fost extrase eșantioanele, în condițiile unui risc de 5%.

12. a) În vederea testării ipotezei de egalitate a salariului mediu cu o valoare fixă, μ_0 , s-a înregistrat un eșantion de volum $n=10$ persoane. În urma prelucrării datelor în SPSS s-au obținut următoarele rezultate:

One-Sample Test

	Test Value = 46					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
VARSTA	-2,637	9	,027	-3,5000	-6,5020	-,4980

Valoarea Sig.=0.027<0,05, se respinge ipoteza de egalitate între mediile comparate. Pe baza datelor din output, se poate considera că există diferențe semnificative între salariul mediu al populației din care a fost extras eșantionul observat și valoarea μ_0 , în condițiile unui risc de 5%.

13. c) În vederea testării ipotezei de egalitate a salariului mediu cu o valoare fixă, μ_0 , s-a înregistrat un eșantion de volum n=10 persoane. În urma prelucrării datelor în SPSS s-au obținut următoarele rezultate:

One-Sample Test						
Test Value = 46						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
VARSTA	-2,637	9	,027	-3,5000	-6,5020	-,4980

În acest caz, se pot formula următoarele ipoteze statistice:

$H_0 : \mu = 46; H_1 : \mu \neq 46$

14. a) În vederea testării ipotezei de egalitate a salariului mediu cu o valoare fixă, μ_0 , s-a înregistrat un eșantion de volum n=10 persoane. În urma prelucrării datelor în SPSS, s-au obținut următoarele rezultate:

One-Sample Test						
Test Value = 22						
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
VARSTA	4,019	9	,003	1,7000	,7432	2,6568

Valoarea Sig.=0.003<0,05, se respinge ipoteza de egalitate între mediile comparate. Pe baza datelor din output, se poate considera că există diferențe semnificative între salariul mediu al populației din care a fost extras eșantionul observat și valoarea μ_0 , în condițiile unui risc de 5%.

15. a) În vederea testării ipotezei de normalitate a unei distribuții s-a înregistrat un eșantion de volum n=40 persoane. În urma prelucrării datelor în SPSS s-au obținut următoarele rezultate:

Test Statistics

	VAR00001
Chi-Square ^a	13,700
df	5
Asymp. Sig.	,018

a. 0 cells (,0%) have expected frequencies less than

5. The minimum expected cell frequency is 6,7.

Valoarea Sig.=0.018<0,05, se respinge ipoteza de normalitate a distribuției, în condițiile unui risc de 5%.

16. c) În vederea testării ipotezei de normalitate a unei distribuții s-a înregistrat un eșantion de volum n=40 persoane. În urma prelucrării datelor în SPSS s-au obținut următoarele rezultate:

One-Sample Kolmogorov-Smirnov Test

		VAR00001
N		40
Normal Parameters ^{a,b}	Mean	6,2500
	Std. Deviation	2,4677
Most Extreme Differences	Absolute	,160
	Positive	,131
	Negative	-,160
Kolmogorov-Smirnov Z		1,010
Asymp. Sig. (2-tailed)		,260

a. Test distribution is Normal.

b. Calculated from data.

Valoarea Sig.=0.260>0,05, se acceptă ipoteza de normalitate a distribuției, în condițiile unui risc de 5%.

11.8. Probleme

11.8.1. Testarea semnificației mediei unei populații

1. În vederea verificării greutății produselor dintr-un lot de piese se extrage aleator simplu repetat un eșantion format din 25 piese. La nivelul eșantionului observat se obține o greutate medie de 340 g și o abatere standard de 10 g. Să se verifice ipoteza potrivit căreia greutatea medie a unei piese din întregul lot concordă cu greutatea medie prevăzută în normele STAS, $\mu_0=330$ g. Riscul asumat este $\alpha=0,05$.

Ipoteze statistice:

$H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$

Statistica test: dacă nu se cunoaște varianța populației și eșantionul este de volum mic, $(n=25)<30$, se folosește statistica t Student, calculată după relația:

$$t_{calc} = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

Regula de decizie:

Dacă $|t_{calc}| > t_{teor} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $|t_{calc}| \leq t_{teor} \Rightarrow$ Se acceptă ipoteza H_0 .

Statistica t Student este:

$$t_{calc} = \frac{340 - 330}{10 / \sqrt{25}} = 5$$

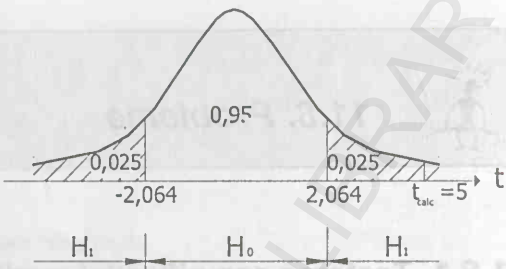
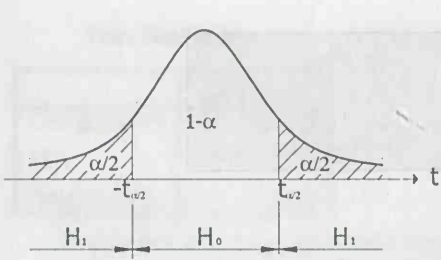
Valoarea teoretică a statisticii test:

Valoarea teoretică a statisticii t Student se citește în Tabela Student pentru un risc $\alpha=0,05$ și $v=n-1=24$ grade de libertate: $t_{0,025;24}=2,064$.

α v	...	$t_{0.025}$	...
...			
24	...	2,064	...
...			

Decizie:

$(t_{calc}=5) > (t_{0,025;24}=2,064)$. În acest caz, se respinge ipoteza H_0 .



Interpretare: Se poate garanta cu o probabilitate de 0,95 că există diferențe semnificative între greutatea medie a unei piese din întregul lot și greutatea prevăzută în normele STAS.

2. În vederea verificării greutății unor produse dintr-un lot de piese se extrage un eșantion format din 100 piese, pentru care se obține o greutate medie de 325 g și o abatere standard de 10g. Să se verifice ipoteza potrivit căreia greutatea medie a unei piese din întregul lot concordă cu greutatea medie prevăzută în normele STAS, $\mu_0=330$ g. Riscul asumat este $\alpha=0,10$.

Ipoteze statistice:

$H_0: \mu = \mu_0$

$H_1: \mu \neq \mu_0$

Statistica test: dacă nu se cunoaște varianța populației și eșantionul este de volum mare, ($n=100$)>30, se folosește statistica Z, calculată după relația:

$$z_{calc} = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

Regula de decizie:

Dacă $|z_{calc}| > z_{teor} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $|z_{calc}| \leq z_{teor} \Rightarrow$ Se acceptă ipoteza H_0 .

Statistica z este:

$$z_{calc} = \frac{325 - 330}{10 / \sqrt{100}} = -5$$

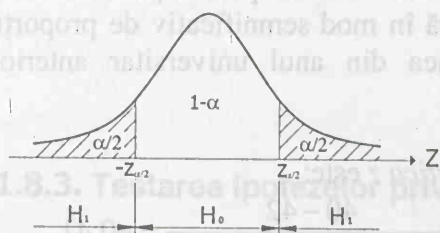
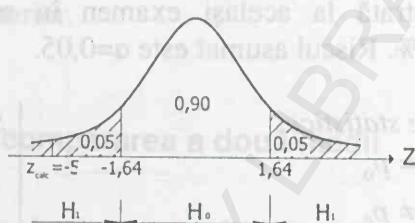
Valoarea teoretică a statisticii test:

Valoarea teoretică a statisticii Z se citește în Tabelul Gauss Laplace pentru un risc $\alpha=0,10$, $z_{teor}=1,64$.

z	... 0.04 ...
⋮	
1.6	... 0,449 ...
⋮	

Decizie:

$(|z_{calc}| = 5) > (z_{teor} = 1,64)$, se respinge

ipoteza H_0 .*Interpretare:*

În aceste condiții, se poate garanta cu o probabilitate de 0,90 că există diferențe semnificative între greutatea medie a unei piese din întregul lot și greutatea medie prevăzută în normele STAS.

3. În vederea testării ipotezei de egalitate a salariului unei populații cu o valoare fixă, μ_0 , s-a înregistrat un eșantion de volum $n=25$ persoane. Să se calculeze numărul gradelor de libertate pentru statistica test care poate fi folosită în acest caz.

Pentru testarea ipotezei de egalitate a mediei unei populații cu o valoare fixă, μ_0 , în cazul eșantioanelor de volum mic ($n < 30$), se folosește statistica test t Student.

Numărul gradelor de libertate corespunzătoare statisticii t Student se calculează după relația:

$$v = n - 1$$

Numărul gradelor de libertate este:

$$v = 25 - 1 = 24 \text{ grade de libertate.}$$

11.8.2. Testarea semnificației unei proporții

1. În vederea aprecierii rezultatelor obținute de ansamblul studenților din anul II de la FEAA la un examen, în sesiunea iulie 2004, se extrage aleator simplu repetat un eșantion format din 100 studenți și se observă că proporția celor care

au luat o notă de până la 5 este de 40%. Să se verifice ipoteza potrivit căreia proporția acestei categorii de studenți diferă în mod semnificativ de proporția înregistrată la același examen în sesiunea din anul universitar anterior, $p_0=42\%$. Riscul asumat este $\alpha=0,05$.

Ipoteze statistice:

$$H_0: p = p_0$$

$$H_1: p \neq p_0$$

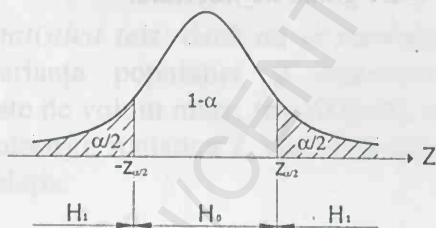
Statistica test: în cazul unui eșan-
tion de volum mare, când nu se
cunoaște varianța populației, se
folosește statistica test z , calculată
după relația:

$$z = \frac{f - p_0}{\sqrt{f \cdot (1 - f) / \sqrt{n}}}$$

Regula de decizie:

Dacă $|z_{\text{calc}}| > z_{\text{teor}} \Rightarrow$ Se respinge ipo-
teza H_0 , se acceptă ipoteza H_1 .

Dacă $|z_{\text{calc}}| \leq z_{\text{teor}} \Rightarrow$ Se acceptă ipo-
teza H_0 .



Statistica z este:

$$z = \frac{40 - 42}{\sqrt{40 \cdot (100 - 40) / \sqrt{100}}} = -0,41$$

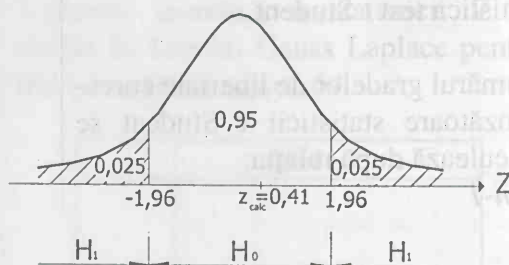
Valoarea teoretică a statisticii test:

Valoarea teoretică a statisticii Z se citește
în Tabelul Gauss Laplace pentru un risc
 $\alpha=0,05$, $z_{\text{teor}}=1,96$.

z	...	0.06	...
$\vdots$			
1.9	...	0,475	...
$\vdots$			

Decizie:

$(|z_{\text{calc}}| = 0,41) < (z_{\text{teor}} = 1,96)$, se acceptă
ipoteza H_0 .



Interpretare:

În aceste condiții, se poate garanta cu o
probabilitate de 0,95 că nu există dife-
rențe semnificative între proporția stu-
denților care au luat o notă de până la 5

în sesiunea iulie 2004 și proporția acestei categorii din anul universitar anterior.

11.8.3. Testarea ipotezelor privind compararea a două medii

1. Pentru două eșantioane extrase aleator simplu repetat de volum $n_1=n_2=100$ persoane s-a înregistrat vârsta și s-au obținut următoarele rezultate: $\bar{x}_1 = 42 \text{ ani}$, $\bar{x}_2 = 37 \text{ ani}$; $s_1 = 4 \text{ ani}$, $s_2 = 7 \text{ ani}$. Să se testeze ipoteza potrivit căreia între vârstele medii ale celor două populații din care au fost extrase eșantioanele observate există diferențe semnificative. Se garantează rezultatul cu o probabilitate de 0,95.

Ipoteze statistice:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

Statistica test: în cazul eșantioanelor de volum mare, pentru testarea ipotezelor statistice se folosește statistica Z, calculată după relația:

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Regula de decizie:

Dacă $|z_{calc}| > z_{teor} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $|z_{calc}| \leq z_{teor} \Rightarrow$ Se acceptă ipoteza H_0 .

În exemplul considerat, $n_1=n_2=100>30$, prin urmare, statistica Z este:

$$z = \frac{42 - 37}{\sqrt{\frac{4^2}{100} + \frac{7^2}{100}}} = 6,2$$

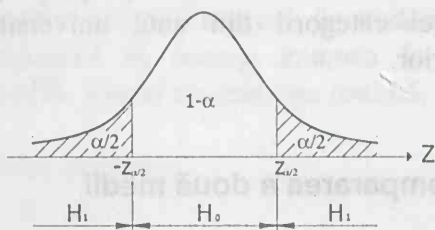
Valoarea teoretică a statisticii test:

Valoarea teoretică a statisticii Z se citește în Tabelul Gauss Laplace pentru un risc $\alpha=0,05$, $z_{teor}=1,96$.

Decizie: $(|z_{calc}| = 6,2) > (z_{teor} = 1,96)$, se respinge ipoteza H_0 .

Interpretare:

Se poate garanta cu o probabilitate de 0,95 că există diferențe semnificative între vârstele medii ale populațiilor din care au fost extrase eșantioanele observate.



2. Se extrag aleator simplu repetat două eșantioane, primul format din 5 persoane de sex masculin, iar al doilea, din 7 persoane de sex feminin. Se înregistrează vârsta acestora și, în urma prelucrării datelor, se obțin următoarele rezultate: $\bar{x}_M = 25 \text{ ani}$, $\bar{x}_F = 26 \text{ ani}$; $s_M^2 = 9$, $s_F^2 = 4$. Să se testeze ipoteza potrivit căreia între vârsta medie a persoanelor de sex masculin și vârsta medie a persoanelor de sex feminin există diferențe semnificative, la nivelul populației din care au fost extrase eșantioanele observate. Se garantează rezultatul cu o probabilitate de 0,95.

Ipoteze statistice:

$$H_0: \mu_M = \mu_F$$

$$H_1: \mu_M \neq \mu_F$$

Statistica test: pentru eșantioane de volum mic, cu varianțe diferite, în testarea ipotezelor statistice se folosește statistica *t* Student:

$$t = \frac{\bar{x}_M - \bar{x}_F}{\sqrt{\frac{s_M^2}{n_1} + \frac{s_F^2}{n_2}}}$$

Regula de decizie:

Dacă $|t_{\text{calc}}| > t_{\text{teor}} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $|t_{\text{calc}}| \leq t_{\text{teor}} \Rightarrow$ Se acceptă ipoteza H_0 .

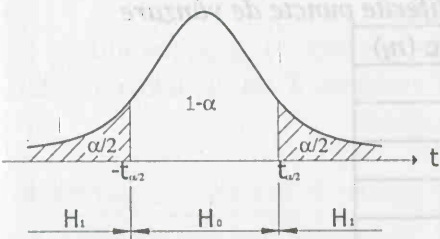
Eșantioanele observate sunt de volum mic: $n_1 = 5$, $n_2 = 7$.

Statistica t Student este:

$$t = \frac{25 - 26}{\sqrt{\frac{9}{5} + \frac{4}{7}}} = -0,65$$

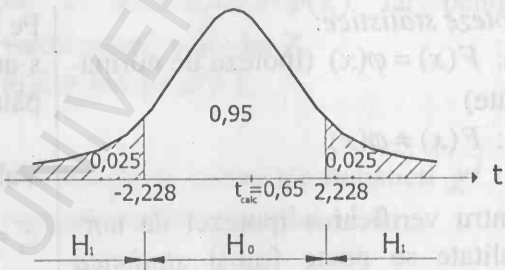
Valoarea teoretică a statisticii test:

Valoarea teoretică a statisticii *t* Student se citește în Tabelul Student pentru un risc $\alpha = 0,05$ și $v = n_1 + n_2 - 2 = 10$ grade de libertate, $t_{0,025,10} = 2,228$.



α	...	$t_{0.025}$	...
v			
...			
10	...	2,228	...
...			

Decizie:
 $(|t_{calc}| = 0,65) < (t_{teor} = 2,228)$, se acceptă ipoteza H_0 .



Interpretare:
Se poate garanta cu o probabilitate de 0,95 că nu există diferențe semnificative între vârsta medie a persoanelor de sex masculin și vârsta medie a persoanelor de sex feminin, la nivelul populației din care au fost extrase eșantioanele observate.

11.8.4. Testarea normalității unei distribuții

- 1. Cererea unui produs a fost observată pe un eșantion de 100 de puncte de vânzare și s-au înregistrat următoarele rezultate:

Tabelul 11.3. Cererea unui produs pentru diferite puncte de vânzare

Cerere (bucăți)	Puncte de vânzare (n_i)
5 - 15	25
15 - 25	30
25 - 35	15
35 - 45	20
45 - 55	10
Total	100

*Date convenționale

Se cere să se testeze și să se interpreteze legea urmată de fenomen, folosind procedeul χ^2 și considerând un risc de 5%.

Ipoteze statistice:

$H_0: F(x) = \varphi(x)$ (ipoteza de normalitate)

$H_1: F(x) \neq \varphi(x)$

Pentru verificarea ipotezei de normalitate se poate folosi *statistica test* χ^2 , care se calculează după relația:

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - n \cdot p_i)^2}{n \cdot p_i},$$

unde:

$$p_i = \varphi(z_i) - \varphi(z_{i-1}), \text{ cu } z_i = \frac{x_i - \bar{x}}{s}.$$

Regula de decizie:

Dacă $\chi_{calc}^2 > \chi_{\alpha, v}^2 \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $\chi_{calc}^2 \leq \chi_{\alpha, v}^2 \Rightarrow$ Se acceptă ipoteza H_0 de normalitate.

Pe baza datelor din tabelul de mai sus, s-au calculat media și abaterea medie pătratică: $\bar{x}=26$ și $s=13,19$.

Valorile z_i se află astfel:

$$z_1 = \frac{10 - 26}{13,19} = -1,21.$$

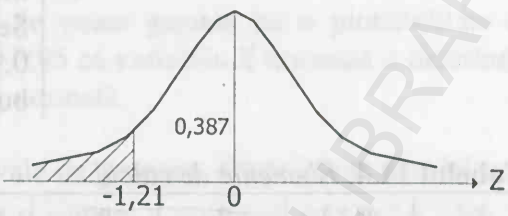
Celelalte valori ale lui z_i sunt prezentate în tabelul 11.4.

Probabilitățile corespunzătoare valorilor z_i se citesc în tabelul Gauss Laplace. Astfel, pentru z_1 aflăm:

$$\varphi(z_i) = \varphi(-1,21) = -\varphi(1,21) = -0,387.$$

z	...	0.01	...
⋮			
1.2	...	0,387	...
⋮			

Valoarea căutată este $P(Z < z_1) = P(Z < -1,21) = 0,5 + \varphi(-1,21) = 0,5 - 0,387 = 0,113$.



Celelalte valori sunt prezentate în tabelul 11.4.

Observație: Pentru valori negative ale lui Z , $P(Z < z_i) = 0,5 - \varphi(z_i)$ iar pentru valori pozitive ale lui Z , $P(Z < z_i) = 0,5 + \varphi(z_i)$.

Elementele de calcul ale statisticii χ^2 sunt prezentate în tabelul 11.4.

Pentru un risc $\alpha = 0,05$ și $v = k - (l + 1) = 5 - (2 + 1) = 2$ grade de libertate, se citește valoarea teoretică a statisticii χ^2 , $\chi^2_{0,05;2} = 5,991$.

$\alpha \backslash v$	...	0,05	...
...			
2	...	5,991	...
...			

Decizia:
 $(\chi^2_{calc} = 27,72) > (\chi^2_{0,05;2} = 5,991)$, se respinge ipoteza H_0 .

Interpretare:

Se poate garanta cu o probabilitate de 0,95 că variabila X nu urmează o distribuție normală.

Tabelul 11.4: Elemente de calcul

z_i	$\varphi(z_i)$	p_i	np_i	$n_i - np_i$	$(n_i - np_i)^2$	$\frac{(n_i - n \cdot p_i)^2}{n \cdot p_i}$
-1,21	0,113	0,113	11,3	13,7	187,69	16,61
-0,45	0,326	0,213	21,3	8,7	75,69	3,55
0,30	0,618	0,292	29,2	-14,2	201,64	6,91
1,06	0,855	0,237	23,7	-3,7	13,69	0,58
1,81	0,964	0,109	10,9	-0,9	0,81	0,07
-	-	0,964	-	-	-	27,72

2. Pentru o variabilă X cu $k=18$ intervale de grupare, observată la nivelul unui eșantion de volum $n=100$, s-a obținut o valoare a statisticii test, $\chi^2 = 15,6$. Se cere să se verifice dacă variabila observată este distribuită normal, considerând un risc $\alpha=0,05$.

Ipoteze statistice:

$H_0: F(x) = \varphi(x)$ (ipoteza de normalitate)

$H_1: F(x) \neq \varphi(x)$

Regula de decizie:

Dacă $\chi^2_{calc} > \chi^2_{\alpha, v} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $\chi^2_{calc} \leq \chi^2_{\alpha, v} \Rightarrow$ Se acceptă ipoteza H_0 de normalitate.

Pentru v grade de libertate și un risc $\alpha=0,05$ se citește în Tabela χ^2 valoarea teoretică, $\chi^2_{0,05;15} = 24,996$.

$\alpha \backslash v$	...	0,05	...
...			
15	...	24,996	...
...			

Decizie:

$(\chi^2_{calc} = 15,6) < (\chi^2_{0,05;15} = 24,996)$, se acceptă ipoteza H_0 .

Interpretare:

Se poate garanta cu o probabilitate de 0,95 că variabila X urmează o distribuție normală.

3. Pentru o variabilă X cu $k=5$ intervale de grupare, observată la nivelul unui eșantion de volum $n=100$, s-a obținut o valoare a statisticii test, $\chi^2 = 25,6$. Se cere să se verifice dacă variabila observată este distribuită normal, considerând un risc $\alpha=0,10$.

Ipoteze statistice:

$H_0: F(x) = \varphi(x)$ (ipoteza de normalitate)

$H_1: F(x) \neq \varphi(x)$

Regula de decizie:

Dacă $\chi^2_{calc} > \chi^2_{\alpha, v} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $\chi^2_{calc} \leq \chi^2_{\alpha, v} \Rightarrow$ Se acceptă ipoteza H_0 de normalitate.

Pentru v grade de libertate și un risc $\alpha=0,10$ se citește în Tabela χ^2 valoarea teoretică, $\chi^2_{0,10;2} = 4,605$.

$\alpha \backslash v$	...	0,10	...
...			
2	...	4,605	...
...			

Decizie: ($\chi^2_{calc} = 25,6$) > ($\chi^2_{0,10;2} = 4,605$)
se respinge ipoteza H_0 .

Interpretare:

Se poate garanta cu o probabilitate de 0,90 că variabila X nu urmează o distribuție normală. Riscul asumat de a respinge pe nedrept ipoteza nulă este de 0,10.

4. La nivelul unei populații de volum $N=100$, o variabilă X urmează o distribuție normală, $X \sim N(15,4)$. Se cere să se calculeze numărul gradelor de libertate ale statisticii χ^2 , considerând că datele sunt grupate pe $k=12$ intervale.

Numărul gradelor de libertate ale statisticii χ^2 se calculează după relația:
 $v=k-(l+1)$,
unde:
 k este numărul intervalelor de grupare;
 l este numărul de parametri estimați.

Numărul gradelor de libertate este:
 $v=k-(l+1)=12-(2+1)=9$ grade de libertate.

α	...	0,05	...
ν	...	...	...
15	...	24,996	...

Capitolul 12

ANALIZA DISPERSIONALĂ (ANOVA)

- ◆ Componente ale variației. Tipuri de ANOVA
- ◆ Componente ale variației și estimatori ai varianțelor în ANOVA unifactorială
- ◆ Condiții. Ipoteze. Regula de decizie în ANOVA unifactorială
- ◆ ANOVA cu SPSS
- ◆ Alegeți răspunsul corect
- ◆ Recunoașteți exemplul corect
- ◆ Răspunsuri corecte
- ◆ Probleme

Analiza dispersională, procedeu statistic cunoscut sub denumirea de ANOVA (*Analysis of Variance*), este utilizat în analiza variației unei variabile în raport cu factorii de influență.

Acest procedeu de analiză statistică constă în descompunerea variației totale a unui ansamblu de date, înregistrate pentru o variabilă X , în componente definite după sursa variației (cauzele acestora) și compararea acestora pentru a stabili dacă factorii considerați cauză au influență semnificativă asupra variabilei X .

12.1. Componente ale variației. Tipuri de ANOVA

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 353-354)

- **Componentele variației** sunt grupate, după sursa de producere a lor, în două categorii:
 - componenta explicată (variație explicată);
 - componenta eroare (reziduu).
- **Exprimarea și măsurarea componentelor variației:**
 - varianța intergrupe (σ_E^2) pentru componenta explicată a variației,
 - varianța intragrupe (σ_R^2) pentru componenta eroare sau reziduu.

Când nu se cunosc varianțele, componentele variației sunt măsurate prin **estimatori ai varianței** și anume:

- estimatorul varianței intergrupe (S_E^2),
- estimatorul varianței intragrupe (S_R^2) și
- estimatorul varianței totale (S_T^2).

• **Tipuri de ANOVA**

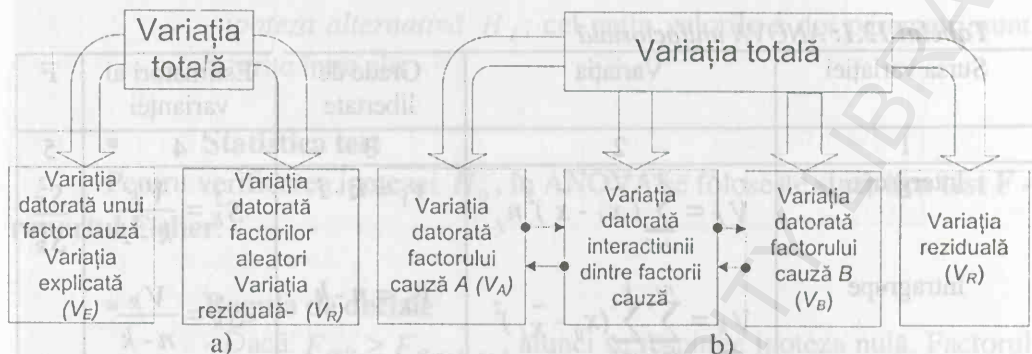


Fig.12.1 (a) ANOVA pentru un factor; (b) ANOVA pentru doi factori

12.2. Componente ale variației și estimatori ai varianțelor în ANOVA unifactorială

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, p. 354)

• **Descompunerea variației totale (V_T):**

- Variația intergrupe (V_E);
- Variația intragrupă (V_R). Relația: $V_T = V_E + V_R$

• **Estimatorii varianței**

- Estimatorul varianței intergrupe: $S_E^2 = \frac{V_E}{k-1}$, unde k reprezintă numărul grupelor (eșantioanelor), iar $k-1$, numărul gradelor de libertate asociat varianței intergrupe.

- Estimatorul varianței intragrupă: $S^2 = \frac{V_R}{n-k}$, unde n reprezintă numărul total al observațiilor.

- Estimatorul corespunzător varianței totale: $S_T^2 = \frac{V_T}{n-1}$.

• **Tabelul de sinteză pentru ANOVA unifactorială**

Tabelul 12.1: ANOVA unifactorială

Sursa variației	Variația	Grade de libertate	Estimatori ai varianței	F
1	2	3	4	5
Intergrupe	$V_E = \sum_{j=1}^k (\bar{x}_j - \bar{x})^2 n_j$	$\nu_1 = k - 1$	$S_E^2 = \frac{V_E}{k - 1}$	$\frac{S_E^2}{S_R^2}$
Intragrupa	$V_R = \sum_{i=1}^{n_j} \sum_{j=1}^k (x_{ij} - \bar{x}_j)^2$	$\nu_2 = n - k$	$S_R^2 = \frac{V_R}{n - k}$	
Totală	$V_T = \sum_{i=1}^n (x_i - \bar{x})^2$	$n - 1$	$S_T^2 = \frac{V_T}{n - 1}$	

Observație! Raportul Fisher – raportul F – este folosit pentru testarea ipotezei H_0 în analiza variației în cazul comparării mediilor a trei sau mai multe eșantioane. În cazul comparării mediilor a două eșantioane se folosește testul Student - testul t.

12.3. Condiții. Ipoteze. Regula de decizie în ANOVA unifactorială

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 355-358)

• **Condiții**

- *Condiția de independență*: eșantioanele aleatoare de efective $n_1, n_2, \dots, n_k$ sunt prelevate din populații într-o manieră independentă;
- *Condiția de normalitate*: distribuțiile populațiilor din care se extrag eșantioanele sunt normale;
- *Condiția de homoscedasticitate*: varianțele celor k populații sunt egale între ele, adică populațiile sunt homoscedastice.

• **Ipoteze**

În analiza variației, considerând un singur factor cauză, se formulează următoarele două ipoteze:

- ipoteza nulă $H_0: \theta_1 = \theta_2 = \dots = \theta_k$, unde θ – parametrul considerat;
- ipoteza alternativă H_1 : cel puțin valorile a doi parametri sunt diferite între ele.

- **Statistica test**

Pentru verificarea ipotezei H_0 , în ANOVA se folosește statistica test F - raportul Fisher.

- **Regula de decizie**

- Dacă $F_{calc} > F_{\alpha, v_1 \text{ si } v_2}$, atunci se respinge ipoteza nulă. Factorul cauză exercită o influență semnificativă asupra variației variabilei considerate.
- Dacă $F_{calc} \leq F_{\alpha, v_1 \text{ si } v_2}$, atunci se acceptă ipoteza nulă, adică se admite egalitatea parametrilor, de exemplu, a mediilor. Factorul cauză nu influențează semnificativ variația variabilei considerate.

12.4. ANOVA cu SPSS

(Vezi Elisabeta Jaba și Ana Grama, *Analiza statistică cu SPSS sub Windows*, Editura Polirom, Iași, 2004, pp. 218-224)

- **Demersul cu SPSS**

- Analyse → Compare Means → One-Way Anova

- **Rezultatele în SPSS**

Rezultatele afișate în output-ul „Oneway” se prezintă astfel:

Oneway

Descriptives

VAR00001

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
rural	9	5,3333	1,0000	,3333	4,5647	6,1020	4,00	7,00
urban	21	7,6190	,9735	,2124	7,1759	8,0622	5,00	9,00
Total	30	6,9333	1,4368	,2623	6,3968	7,4698	4,00	9,00

ANOVA

VAR00001

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	32,914	1	32,914	34,194	,000
Within Groups	26,952	28	,963		
Total	59,867	29			



12.5. Alegeți răspunsul corect

- Componentele variației sunt
 - variația intergrupe
 - variația totală
 - variația intragrupe
 - variația parțială
- Condițiile cerute de ANOVA sunt
 - independența eșantioanelor
 - normalitatea distribuțiilor populațiilor
 - homoscedasticitatea populațiilor
 - heterodasticitatea populațiilor
- Pentru testarea egalității a trei sau mai multor medii se folosește
 - testul F
 - testul t
 - testul Z
- Testul F este un raport între
 - estimatorul varianței reziduale și al celei totale
 - estimatorul varianței explicate și al celei totale
 - estimatorul varianței explicate și estimatorul varianței reziduale
- Variația intergrupe măsoară
 - variația mediilor grupelor față de media lor
 - variația din interiorul fiecărei grupe
 - variația totală
- Variația intragrupe măsoară
 - variația mediilor grupelor față de media lor
 - variația din interiorul fiecărei grupe
 - variația totală
- Variația intergrupe se calculează după relația:

$$a. V_E = \sum_{j=1}^k (\bar{x}_j - \bar{x})^2 \cdot n_j$$

$$b. V_R = \sum_{i=1}^{n_j} \sum_{j=1}^k (x_{ij} - \bar{x}_j)^2$$

$$c. \quad V_T = \sum_{i=1}^n (x_i - \bar{x})^2$$

8. Variația intragrupe se calculează după relația:

$$a. \quad V_E = \sum_{j=1}^k (\bar{x}_j - \bar{x})^2 \cdot n_j$$

$$b. \quad V_R = \sum_{i=1}^{n_j} \sum_{j=1}^k (x_{ij} - \bar{x}_j)^2$$

$$c. \quad V_T = \sum_{i=1}^n (x_i - \bar{x})^2$$

9. Variația totală se calculează după relația:

$$a. \quad V_E = \sum_{j=1}^k (\bar{x}_j - \bar{x})^2 \cdot n_j$$

$$b. \quad V_R = \sum_{i=1}^{n_j} \sum_{j=1}^k (x_{ij} - \bar{x}_j)^2$$

$$c. \quad V_T = \sum_{i=1}^n (x_i - \bar{x})^2$$

10. Estimatorul varianței intergrupe se calculează după relația:

$$a. \quad S_E^2 = \frac{V_E}{k-1}$$

$$b. \quad S_R^2 = \frac{V_R}{n-k}$$

$$c. \quad S_T^2 = \frac{V_T}{n-1}$$

11. Estimatorul varianței intragrupe se calculează după relația:

$$a. \quad S_E^2 = \frac{V_E}{k-1}$$

$$b. \quad S_R^2 = \frac{V_R}{n-k}$$

$$c. \quad S_T^2 = \frac{V_T}{n-1}$$

12. Estimatorul varianței totale se calculează după relația:

$$a. \quad S_E^2 = \frac{V_E}{k-1}$$

b. $S_R^2 = \frac{V_R}{n-k}$

c. $S_T^2 = \frac{V_T}{n-1}$

13. Numărul gradelor de libertate asociat varianței intergrupe este

a. $v_1 = k - 1$

b. $v_2 = n - k$

c. $v = n - 1$

14. Numărul gradelor de libertate asociat varianței intragrupe este

a. $v_1 = k - 1$

b. $v_2 = n - k$

c. $v = n - 1$

15. Numărul gradelor de libertate asociat varianței totale este

a. $v_1 = k - 1$

b. $v_2 = n - k$

c. $v = n - 1$

ANOVA (Vârsta respondent)	Suma pătratelor	Libertate	Medie pătrată	F	Probabilitate
Intergrupe	200.00	2	100.00	1.00	0.37
Intragrupe	180.00	18	10.00		
Total	380.00	20			



12.6. Recunoașteți exemplul corect

1. În urma unei analize ANOVA s-a obținut o valoare calculată a raportului $F=7,54$. Dacă valoarea teoretică a statisticii F este $F_{0,05;2;14}=3.739$, atunci se poate considera că

- se garantează cu o probabilitate de 0,95 că factorul de grupare are o influență semnificativă asupra variației variabilei considerate
- se garantează cu o probabilitate de 0,05 că factorul de grupare are o influență semnificativă asupra variației variabilei considerate
- se garantează cu o probabilitate de 0,95 că factorul de grupare nu are o influență semnificativă asupra variației variabilei considerate

2. S-a înregistrat rata șomajului asupra unui eșantion de 972 persoane, la nivelul regiunilor unei țări, în anul 2004. S-a testat, prin metoda ANOVA, ipoteza potrivit căreia factorul de grupare „Regiunea” are o influență semnificativă asupra ratei șomajului. S-au obținut următoarele rezultate:

ANOVA Unemployed and Looking for Work a Month

	Suma pătraterelor abaterilor	Grade de libertate	Estimatori ai varianței	F	Sig.
Intergrupe	6,302E-02	2	3,151E-02	,552	,576
Intragrupa	55,359	970	5,707E-02		
Total	55,422	972			

Ipotezele statistice care se pot formula în acest caz sunt:

a. $H_0: \mu_1 = \mu_2; H_1: \mu_1 \neq \mu_2$

$\Rightarrow$ b. $H_0: \mu_1 = \mu_2 = \mu_3; H_1$: rata medie a șomajului din cel puțin două regiuni diferă

* c. $H_0: \mu_1 = \mu_2 = \dots = \mu_{970}; H_1$: rata medie a șomajului din cel puțin două regiuni diferă

3. S-a înregistrat rata șomajului asupra unui eșantion de 972 persoane, la nivelul regiunilor unei țări, în anul 2004. S-a testat, prin metoda ANOVA,

ipoteza potrivit căreia factorul de grupare „Regiunea“ are o influență semnificativă asupra ratei șomajului. S-au obținut următoarele rezultate:

ANOVA Unemployed and Looking for Work a Month

	Suma pătratelor abaterilor	Grade de libertate	Estimatori ai varianței	F	Sig.
Intergrupe	6,302E-02	2	3,151E-02	,552	,576
Intragrupe	55,359	970	5,707E-02		
Total	55,422	972			

În această situație, se poate considera că

- a. se garantează cu o probabilitate de 0,95 că factorul de grupare are o influență semnificativă asupra variației variabilei considerate
- b. se garantează cu o probabilitate de 0,05 că factorul de grupare are o influență semnificativă asupra variației variabilei considerate
- c. se garantează cu o probabilitate de 0,95 că factorul de grupare nu are o influență semnificativă asupra variației variabilei considerate

4. Pentru a verifica dacă structura după vârstă a populației diferă de la o regiune la alta s-a înregistrat această caracteristică pentru un eșantion de 1514 persoane, la nivelul regiunilor unei țări, în anul 2004. Aplicând metoda ANOVA, cu ajutorul SPSS, s-au obținut următoarele rezultate:

ANOVA (Vârsta respondenților)

	Suma pătratelor abaterilor	Grade de libertate	Estimatori ai varianței	F	Sig.
Intergrupe	3300,022	2	1650,011	5,232	,005
Intragrupe	476532,382	1511	315,376		
Total	479832,404	1513			

În această situație, se poate considera că

- a. se garantează cu o probabilitate de 0,95 că factorul de grupare are o influență semnificativă asupra variației variabilei considerate
- b. se garantează cu o probabilitate de 0,05 că factorul de grupare are o influență semnificativă asupra variației variabilei considerate
- c. se garantează cu o probabilitate de 0,95 că factorul de grupare nu are o influență semnificativă asupra variației variabilei considerate



12.7. Răspunsuri corecte

12.5. Alegeți răspunsul corect

1. a) și c) Componentele variației sunt variația intergrupe și variația intragrupe.
2. a), b) și c) Condițiile cerute de ANOVA sunt independența eșantioanelor, normalitatea distribuțiilor populațiilor și homoscedasticitatea populațiilor.
3. a) Pentru testarea egalității a trei sau mai multor medii se folosește testul F .
4. c) Testul F este un raport între estimatorul varianței explicate și estimatorul varianței reziduale.
5. a) Variația intergrupe măsoară variația mediilor grupelor față de media lor.
6. b) Variația intragrupe măsoară variația din interiorul fiecărei grupe.
7. a) Variația intergrupe se calculează după relația: $V_E = \sum_{j=1}^k (\bar{x}_j - \bar{x})^2 \cdot n_j$.
8. b) Variația intragrupe se calculează după relația: $V_R = \sum_{i=1}^{n_j} \sum_{j=1}^k (x_{ij} - \bar{x}_j)^2$.
9. c) Variația totală se calculează după relația: $V_T = \sum_{i=1}^n (x_i - \bar{x})^2$.
10. a) Estimatorul varianței intergrupe se calculează după relația: $S_E^2 = \frac{V_E}{k-1}$.
11. b) Estimatorul varianței intragrupe se calculează după relația: $S_R^2 = \frac{V_R}{n-k}$.
12. c) Estimatorul varianței totale se calculează după relația: $S_T^2 = \frac{V_T}{n-1}$.
13. a) Numărul gradelor de libertate asociat varianței intergrupe este $v_1 = k-1$.
14. b) Numărul gradelor de libertate asociat varianței intragrupe este $v_2 = n-k$.
15. c) Numărul gradelor de libertate asociat varianței totale este $\hat{v} = n-1$.

12.6. Recunoașteți exemplul corect

1. a) În urma unei analize ANOVA s-a obținut o valoare calculată a raportului $F=7,54$. Dacă valoarea teoretică a statisticii F este $F_{0,05,2,14}=3.739$, atunci se poate considera că se garantează cu o probabilitate de 0,95 că factorul de grupare are o influență semnificativă asupra variației variabilei considerate.

2. b) Deoarece $v_1 = k - 1 = 2$, numărul regiunilor pentru care s-a înregistrat rata șomajului este $k=3$. Ipotezele statistice care se pot formula în acest caz sunt:

$$H_0: \mu_1 = \mu_2 = \mu_3;$$

H_1 : rata medie a șomajului din cel puțin două regiuni diferă.

3. c) Deoarece $Sig. = 0,576 > \alpha = 0,05$, se acceptă ipoteza H_0 . Se poate garanta cu o probabilitate de 0,95 că factorul de grupare „Regiunea” nu are o influență semnificativă asupra ratei șomajului.

4. a) Deoarece $Sig. = 0,005 < \alpha = 0,05$, se respinge ipoteza H_0 . Se poate garanta cu o probabilitate de 0,95 că factorul de grupare „Regiunea” are o influență semnificativă asupra ratei șomajului.

	V_1	V_2	V_3	V_4
V_1	...	...	...	...
V_2	...	...	...	...
V_3	...	...	...	...
V_4	...	...	...	...



12.8. Probleme

1. S-a înregistrat câștigul salarial nominal net asupra unui eșantion de 5000 persoane la nivelul celor 8 regiuni ale României, în anul 2003. Aplicând metoda ANOVA, s-au obținut următoarele rezultate: $V_E=630$, $V_R=3400$. Se cere să se testeze ipoteza potrivit căreia factorul de grupare „Regiunea” are o influență semnificativă asupra valorii câștigului salarial nominal net. Considerăm un risc $\alpha=0,05$.

Pentru variabila considerată sunt îndeplinite condițiile cerute de ANOVA (condiția de independență, de normalitate și de homoscedasticitate).

Ipoteze statistice:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_8$$

H_1 : câștigul salarial mediu din cel puțin două regiuni diferă

Statistica test:

$$F = \frac{S_E^2}{S_R^2},$$

unde:

$$S_E^2 = \frac{V_E}{v_1} = \frac{V_E}{k-1},$$

$$S_R^2 = \frac{V_R}{v_2} = \frac{V_R}{n-k},$$

Regula de decizie:

Dacă $F_{calc} > F_{\alpha; v_1; v_2} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $F_{calc} \leq F_{\alpha; v_1; v_2} \Rightarrow$ Se acceptă ipoteza H_0 .

Estimatorii varianțelor:

$$S_E^2 = \frac{V_E}{k-1} = \frac{630}{8-1} = 90;$$

$$S_R^2 = \frac{V_R}{n-k} = \frac{3400}{5000-8} = 0,681.$$

Statistica test:

$$F = \frac{S_E^2}{S_R^2} = \frac{90}{0,681} = 132,16$$

Valoarea teoretică:

Valoarea teoretică a statisticii F se citește în Tabela Fisher pentru un risc $\alpha=0,05$ și $v_1=k-1=8-1=7$, $v_2=n-k=5000-8=4992$ grade de libertate. Aceasta este $F_{0,05;7;4992}=2,01$.

- pentru $\alpha=0,05$:

$v_1 \backslash v_2$	...	7	...
...			
7	...	2,01	...

Decizie:

Deoarece $(F_{calc} = 13216) > (F_{\alpha; v_1; v_2} = 2,01)$, se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Interpretare:

Se poate garanta cu o probabilitate de 0,95 că factorul de grupare „Regiunea” are o influență semnificativă asupra valorii câștigului salarial nominal net.

2. S-a înregistrat câștigul salarial nominal net asupra unui eșantion de 5000 persoane la nivelul celor 8 regiuni ale României, în anul 2003. Aplicând metoda ANOVA, s-au obținut următoarele rezultate: $S_E^2 = 300, V_R = 3400$. Se cere să se testeze ipoteza potrivit căreia factorul de grupare „Regiunea” are o influență semnificativă asupra valorii câștigului salarial nominal net ($\alpha=0,01$).

Pentru variabila considerată sunt îndeplinite condițiile cerute de ANOVA (condiția de independență, de normalitate și de homoscedasticitate).

Ipoteze statistice:

$H_0: \mu_1 = \mu_2 = \dots = \mu_8$

H_1 : câștigul salarial mediu din cel puțin două regiuni diferă

Statistica test:

$$F = \frac{S_E^2}{S_R^2},$$

unde:

$$S_E^2 = \frac{V_E}{v_1} = \frac{V_E}{k-1},$$

$$S_R^2 = \frac{V_R}{v_1} = \frac{V_R}{k-1},$$

Estimatorul varianței reziduale:

$$S_R^2 = \frac{V_R}{n-k} = \frac{3400}{5000-8} = 0,681.$$

Statistica test:

$$F = \frac{S_E^2}{S_R^2} = \frac{300}{0,681} = 440,53$$

Valoarea teoretică:

Valoarea teoretică a statisticii F se citește în Tabela Fisher pentru un risc $\alpha=0,01$ și $v_1=k-1=8-1=7$, $v_2=n-k=5000-8=4992$ grade de libertate, $F_{0,01;7;4992}=2,639$.

- pentru $\alpha=0,01$:

Regula de decizie:

Dacă $F_{calc} > F_{\alpha;v_1;v_2} \Rightarrow$ Se respinge

ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $F_{calc} \leq F_{\alpha;v_1;v_2} \Rightarrow$ Se acceptă ipoteza H_0 .

$v_1 \backslash v_2$	...	7	...
...			
...			
À	...	2,639	...

Decizia:

Deoarece

$(F_{calc} = 440,53) > (F_{\alpha;v_1;v_2} = 2,639)$, se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Interpretare:

Se poate garanta cu o probabilitate de 0,99 că factorul de grupare „Regiunea” are o influență semnificativă asupra valorii câștigului salarial nominal net.

3. Aplicând metoda ANOVA, cu ajutorul SPSS, s-au obținut următoarele rezultate:

ANOVA

	Suma pătratelor abaterilor	Grade de libertate	Estimatori ai varianței	F	Sig.
Intergrupe	3300,022	2	1650,011	5,232	,005
Intragrupa	476532,382	1511	315,376		
Total	479832,404	1513			

Se cere să se interpreteze rezultatele din tabelul ANOVA. Variabila numerică este vârsta populației, iar variabila de grupare este regiunea. Considerăm un risc $\alpha=0,05$.

Ipoteze statistice:

$H_0: \mu_1 = \mu_2 = \mu_3$

H_1 : vârsta medie din cel puțin două

Raportul F este:

$$F = \frac{S_E^2}{S_R^2} = \frac{1650,011}{315,376} = 5,232$$

regiuni diferă

Statistica test:

$$F = \frac{S_E^2}{S_R^2}$$

Regula de decizie:

Dacă $F_{calc} > F_{\alpha;v_1;v_2}$ sau $Sig < \alpha \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $F_{calc} \leq F_{\alpha;v_1;v_2}$ sau $Sig > \alpha \Rightarrow$ Se acceptă ipoteza H_0 .

Valoarea teoretică:

Valoarea teoretică a statisticii F se citește în Tabela Fisher pentru un risc $\alpha=0,05$ și $v_1=2$, $v_2=1511$ grade de libertate, $F_{0,05;2;1511}=2,996$.

- pentru $\alpha=0,05$:

$v_1 \backslash v_2$	...	2	...
...			
2	...	2,996	...

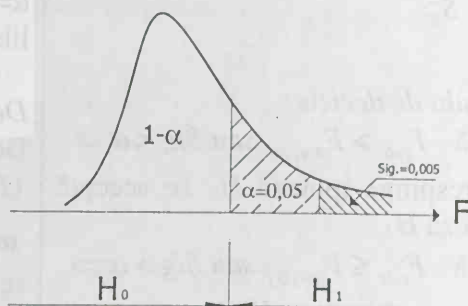
Decizia:

Deoarece

$$(F_{calc} = 5,232) > (F_{\alpha;v_1;v_2} = 2,996)$$

$$\text{sau } (Sig. = 0,005) < (\alpha = 0,05)$$

se respinge ipoteza H_0 , se acceptă ipoteza H_1 .



Interpretare:

Se poate garanta cu o probabilitate de 0,95 că factorul de grupare „Regiunea” are o influență semnificativă asupra vârstei persoanelor.

4. Aplicând metoda ANOVA, cu ajutorul SPSS, s-au obținut următoarele rezultate:

ANOVA

	Suma pătratelor abaterilor	Grade de libertate	Estimatori ai varianței	F	Sig.
Intergrupe	6,302E-02	2	3,151E-02	,552	,576
Intragrupa	55,359	970	5,707E-02		
Total	55,422	972			

Se cere să se interpreteze rezultatele din tabelul ANOVA. Variabila numerică este rata șomajului, iar variabila de grupare este regiunea. Considerăm un risc $\alpha=0,05$.

Ipoteze statistice:

$$H_0: \mu_1 = \mu_2 = \mu_3$$

H_1 : rata medie a șomajului din cel puțin două regiuni diferă

Statistica test:

$$F = \frac{S_E^2}{S_R^2}$$

Regula de decizie:

Dacă $F_{calc} > F_{\alpha;v_1;v_2}$ sau $Sig < \alpha \Rightarrow$

Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $F_{calc} \leq F_{\alpha;v_1;v_2}$ sau $Sig > \alpha \Rightarrow$

Se acceptă ipoteza H_0 .

Raportul F este:

$$F = \frac{S_E^2}{S_R^2} = \frac{0,03151}{0,05707} = 0,552$$

Valoarea teoretică:

Valoarea teoretică a statisticii F se citește în Tabela Fisher pentru un risc $\alpha=0,05$ și $v_1=2$, $v_2=970$ grade de libertate, $F_{0,05;2;970}=2,996$.

Decizie:

Deoarece

$$(F_{calc} = 0,552) < (F_{\alpha;v_1;v_2} = 2,996)$$

$$\text{sau } (Sig. = 0,576) > (\alpha = 0,05),$$

se acceptă ipoteza H_0 .

Interpretare:

Se poate garanta cu o probabilitate de 0,95 că factorul de grupare „Regiunea” nu are o influență semnificativă asupra ratei șomajului.

Capitolul 13

ANALIZA DE REGRESIE ȘI CORELAȚIE

- ◆ Probleme ale analizei de regresie și corelație
- ◆ Metode de analiză a legăturilor statistice
- ◆ Modele de regresie cu două variabile
- ◆ Măsurarea intensității legăturii dintre două variabile
- ◆ Corelația neparametrică
- ◆ Analiza de regresie și corelație cu SPSS
- ◆ Alegeți răspunsul corect
- ◆ Recunoașteți exemplul corect
- ◆ Răspunsuri corecte
- ◆ Probleme

13.1. Probleme ale analizei de regresie și corelație

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economică, București, 2002, pp. 371-374)

- **Noțiuni**

Regresia exprimă o legătură de tip statistic și anume regresia în medie² cu privire la comportamentul unor variabile.

Corelația exprimă raporturi reciproce între anumite caracteristici.

Analiza de regresie este folosită pentru:

- estimarea valorilor unei variabile, considerând valorile altei/altor variabile;
- evaluarea măsurii în care variabila dependentă poate fi explicată prin variabila independentă, sau printr-un set de variabile independente;
- identificarea unui subset din mai multe variabile independente care trebuie luat în calcul pentru estimarea variabilei dependente.

- **Model de regresie**

- În expresie generală, un model de regresie poate fi scris:

$$Y = f(X_1, X_2, \dots, X_n) + \varepsilon,$$

în care:

Y – variabila dependentă (rezultativă), aleatoare;

$X_1, \dots, X_n$ – variabile independente (factoriale) nonaleatoare;

ε – variabila aleatoare eroare sau reziduu.

- *modelul de regresie simplă liniară* se poate scrie:

$$Y = \alpha + \beta X + \varepsilon,$$

² În literatura de specialitate, expresia **regresie în medie** sau cum a scăpat omenirea de „gigantism” și „piticism” este legată de cercetările lui Fr. Galton asupra eredității. Galton, observând modul în care evoluează înălțimea copiilor față de cea a părinților, a ajuns la concluzia că, de regulă, din părinți de talie mare se nasc copii cu o talie inferioară lor, iar din părinți de talie redusă se nasc copii cu talie mai mare decât a părinților. Dacă din părinți de talie mare s-ar fi născut copii cu talie și mai mare ș.a.m.d. s-ar fi ajuns la gigantism, sau invers, la piticism. Se înțelege că legitatea descoperită, ca orice legitate statistică, se verifică nu pe cazuri izolate, ci la nivelul colectivităților de volum mare. (Fr. Galton, *Natural Inheritance*, Macmillan, London, 1889).

- *modelul de regresie multiplu liniar*, cu două sau mai multe variabile factoriale poate fi scris:

$$Y = \alpha + \beta_1 X_1 + \dots + \beta_n X_n + \varepsilon.$$

13.2. Metode de analiză a legăturilor statistice

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 374-378)

- Metode elementare
- Metoda grupărilor statistice
- Metoda tabelului de corelație
- Metoda grafică
- Metode analitice

13.3. Modele de regresie cu două variabile

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 378-390)

- **Demersul analizei de regresie**
 - *Construirea corelogramei*, respectiv a *norului de puncte* (*diagrama de dispersie sau scatterplot*);
 - *Aproximarea formei legăturii* printr-un model de regresie și scrierea ecuației corespunzătoare;
 - *Estimarea parametrilor ecuației de regresie* (pe baza metodei celor mai mici pătrate) și interpretarea regresiei în funcție de semnul și valoarea lor;
 - *Testarea semnificației parametrilor de regresie*.

- **Parametrii ecuației de regresie**

Pentru un model de regresie simplu liniar de forma: $Y = \alpha + \beta X + \varepsilon$, parametrii sunt:

α - *ordonata la origine* arată valoarea variabilei Y când $X = 0$;

β - *panta dreptei*, numit și coeficient de regresie.

Semnul parametrului de regresiei β indică direcția legăturii dintre cele două variabile corelate:

$\beta > 0$ legătură directă (pozitivă);

$\beta = 0$ nu există legătură;

$\beta < 0$ legătură inversă (negativă).

Parametrul de regresie β arată gradul de dependență dintre variabile, respectiv cu cât crește sau scade Y la o creștere sau la o scădere a variabilei X cu o unitate.

13.3.1. Estimarea punctuală a parametrilor ecuației de regresie

Pentru estimarea parametrilor ecuației de regresie, în practică, se utilizează, de regulă, metoda celor mai mici pătrate (MCMMP):

$$S = \sum e_i^2 = \sum (y_i - y_{x_i})^2 = \text{minim},$$

unde:

y_i reprezintă valorile observate;

y_{x_i} reprezintă o valoare estimată a ecuației de regresie la nivelul eșantionului, $y_{x_i} = a + bx_i$. Valorile a și b sunt estimații ale parametrilor α și β ai ecuației de regresie.

Aplicând metoda determinanților, se obțin pentru a și b următoarele relații de calcul:

$$b = \frac{\Delta b}{\Delta} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}, \quad i = \overline{1, n}$$

$$a = \frac{\Delta a}{\Delta} = \frac{\sum y_i \sum x_i^2 - \sum x_i \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

$$a = \bar{y} - b\bar{x}$$

13.3.2. Estimarea prin interval de încredere

Intervalul de încredere pentru coeficienții de regresie β și α estimat pentru un eșantion observat este definit de relațiile:

$$(b \pm t_{\alpha/2} \cdot s_{\hat{\beta}})$$

$$(a \pm t_{\alpha/2} \cdot s_{\hat{\alpha}}),$$

unde: $s_{\hat{\beta}} = \sqrt{\frac{s_e^2}{\sum_i (x_i - \bar{x})^2}}$

$$s_{\hat{\alpha}} = \sqrt{\frac{\sum_i x_i^2}{n \sum_i (x_i - \bar{x})^2} \cdot s_e^2}$$

$$s_e = \sqrt{\frac{\sum_i e_i^2}{n-2}} = \sqrt{\frac{\sum_i (y_i - a - bx_i)^2}{n-2}}$$

13.4. Măsurarea intensității legăturii dintre două variabile

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 390-397)

13.4.1. Coeficientul de corelație

- Estimarea coeficientului de corelație

Valoarea coeficientului de corelație este cuprinsă între -1 și +1:

$$-1 \leq \rho \leq +1$$

La nivelul unui eșantion de volum n , coeficientul de corelație ρ este estimat prin *coeficientul de corelație empiric* propus de K. Pearson, r_{yx} , calculat după relația:

$$r_{yx} = \frac{\text{cov}(x, y)}{s_x \cdot s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n \cdot s_x \cdot s_y}$$

Dezvoltând relația de mai sus, se obține o *formulă de calcul simplificat al coeficientului de corelație empiric* bazată pe elementele calculate deja cu ocazia calculării coeficientului de regresie:

$$r_{yx} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2][n \sum y_i^2 - (\sum y_i)^2]}}, \quad i = \overline{1, n}$$

- **Testarea coeficientului de corelație**

- *Testul t (Student)*, folosit pentru verificarea semnificației coeficientului de corelație simplă, este:

$$t = \frac{\hat{\rho}}{\hat{\sigma}_{\hat{\rho}}} = \frac{\hat{\rho}\sqrt{n-2}}{\sqrt{1-\hat{\rho}^2}},$$

unde:

t este o statistică Student cu $(n-2)$ grade de libertate.

$\hat{\sigma}_{\hat{\rho}}$ este estimatorul abaterii medii pătratice a lui $\hat{\rho}$ (estimatorul lui ρ):

$$\hat{\sigma}_{\hat{\rho}} = \sqrt{\frac{1-\hat{\rho}^2}{n-2}}$$

La nivelul unui eșantion observat, se determină valoarea calculată a testului:

$$t = \frac{r_{yx} \cdot \sqrt{n-2}}{\sqrt{1-r_{yx}^2}}$$

- *Decizia:*

Valoarea calculată a lui t se compară cu valoarea teoretică obținută din tabelul t , pentru $n-2$ grade de libertate și pentru nivelul de semnificație stabilit. Dacă $t_{calc.} > t_{tab.}$ atunci se respinge H_0 și se trage concluzia că între variabilele cercetate există o legătură semnificativă, deci coeficientul de corelație este semnificativ statistic.

13.4.2. Raportul de corelație

- **Estimarea raportului de corelație**

Valoarea raportului de corelație este un număr cuprins în intervalul: $0 \leq \eta \leq 1$.

La nivelul unui eșantion observat, raportul de corelație η se poate estima pe baza valorilor observate, după relația:

$$\eta_{y_x} = \sqrt{\frac{a \sum y_i + b \sum x_i y_i - \frac{1}{n} (\sum y_i)^2}{\sum y_i^2 - \frac{1}{n} (\sum y_i)^2}}, \quad i = \overline{1, n}$$

Valoarea la pătrat a raportului de corelație reprezintă **raportul de determinație**.

• **Testarea raportului de corelație**

- Testul F : $F = \frac{n-k}{k-1} \cdot \frac{\eta^2}{1-\eta^2}$,

unde:

n reprezintă numărul valorilor observate;

k – numărul parametrilor estimați ai modelului de regresie;

$\hat{\eta}$ – estimatorul raportului de corelație.

Statistica F urmează o lege de distribuție Snedecor-Fisher de $\nu_1 = k-1$ și $\nu_2 = n-k$ grade de libertate.

La nivelul unui eșantion observat, se determină valoarea calculată a testului:

$$F = \frac{n-k}{k-1} \cdot \frac{\eta^2}{1-\eta^2}$$

- Decizia:

Dacă $F_{calc.} > F_{tab.}$ cu $\nu_1 = k-1$ și $\nu_2 = n-k$ grade de libertate, atunci se trage concluzia că variabila factorială influențează semnificativ comportarea variabilei rezultative, deci raportul de corelație este semnificativ statistic.

13.5. Corelația neparametrică

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economică, București, 2002, pp. 403-406)

- Coeficientul de concordanță Fechner
- Coeficienții de corelație a rangurilor (coeficientul Spearman și Kendall)
- Coeficientul de asociere

13.6. Analiza de regresie și corelație cu SPSS

(Vezi Elisabeta Jaba și Ana Grama, *Analiza statistică cu SPSS sub Windows*, Editura Polirom, Iași, 2004, pp. 240-257)

- Demersul în SPSS
 - Analyse → Correlate → Bivariate;
 - Analyse → Regression → Linear

- Rezultatele în SPSS

Rezultatele afișate în fereastra „Correlations” se prezintă astfel:

Correlations

Correlations

		VAR00001	VAR00003
VAR00001	Pearson Correlation	1,000	,943**
	Sig. (2-tailed)		,000
	N	10	10
VAR00003	Pearson Correlation	,943**	1,000
	Sig. (2-tailed)	,000	
	N	10	10

** Correlation is significant at the 0.01 level (2-tailed).

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	VAR00003 ^a		Enter

- a. All requested variables entered.
- b. Dependent Variable: VAR00001

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,943 ^a	,889	,875	,7102

- a. Predictors: (Constant), VAR00003
- b. Dependent Variable: VAR00001

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	32,365	1	32,365	64,162	,000 ^a
	Residual	4,035	8	,504		
	Total	36,400	9			

- a. Predictors: (Constant), VAR00003
- b. Dependent Variable: VAR00001

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,253	,800		,316	,760
	VAR00003	,904	,113	,943	8,010	,000

- a. Dependent Variable: VAR00001

Nonparametric Correlations

Correlations

			VAR00001	VAR00003
Kendall's tau_b	VAR00001	Correlation Coefficient	1,000	,850**
		Sig. (2-tailed)	,	,001
		N	10	10
	VAR00003	Correlation Coefficient	,850**	1,000
		Sig. (2-tailed)	,001	,
		N	10	10
Spearman's rho	VAR00001	Correlation Coefficient	1,000	,914**
		Sig. (2-tailed)	,	,000
		N	10	10
	VAR00003	Correlation Coefficient	,914**	1,000
		Sig. (2-tailed)	,000	,
		N	10	10

** Correlation is significant at the .01 level (2-tailed).



13.7. Alegeți răspunsul corect

1. Analiza de regresie este o metodă statistică care permite
 - a. studierea și măsurarea relației care există între două sau mai multe variabile
 - b. descoperirea legii relative la forma legăturilor dintre variabile
 - c. măsurarea intensității legăturii dintre variabile
2. Analiza de corelație este o metodă care permite
 - a. studierea și măsurarea relației care există între două sau mai multe variabile
 - b. descoperirea legii relative la forma legăturilor dintre variabile
 - c. măsurarea intensității legăturii dintre variabile
3. Variabila factorială (independentă) este
 - a. variabila care explică variația altei variabile
 - b. variabila estimată, Y
 - c. variabila care sintetizează toți factorii care influențează variabila Y dar sunt nespecificați în model
4. Variabila rezultativă (dependentă) este
 - a. variabila care explică variația altei variabile
 - b. variabila estimată, Y
 - c. variabila care sintetizează toți factorii care influențează variabila Y dar sunt nespecificați în model
5. Variabila reziduală (aleatoare) este
 - a. variabila care explică variația altei variabile
 - b. variabila estimată, Y
 - c. variabila care sintetizează toți factorii care influențează variabila Y dar sunt nespecificați în model
6. După numărul variabilelor corelate, legăturile statistice pot fi
 - a. legături simple
 - b. legături multiple
 - c. legături directe
 - d. legături inverse
7. După sensul legăturii statistice, acestea pot fi
 - a. legături simple
 - b. legături multiple

- c. legături directe
 - d. legături inverse
8. După formă, legăturile statistice pot fi
- a. legături simple
 - b. legături rectilinii
 - c. legături curbilinii
9. Metodele elementare de analiză a legăturilor statistice sunt
- a. metoda seriilor statistice paralele interdependente
 - b. metoda grupărilor statistice
 - c. metoda tabelului de corelație
 - d. metoda grafică
10. Metoda celor mai mici pătrate minimizează:
- a. suma pătratelor erorilor
 - b. produsul abaterilor de la medie
 - c. suma abaterilor de la medie
11. Intervalul de încredere pentru parametrul β este:
- a. probabilist
 - b. determinist
 - c. de lungime minimă
12. Intervalul de încredere pentru parametrul β , estimat pentru un eșantion observat, este definit de relația:
- a. $(b \pm t_{\alpha/2} \cdot s_{\hat{\beta}})$
 - b. $(a \pm t_{\alpha/2} \cdot s_{\hat{a}})$
 - c. $(b \pm t_{\alpha/2} \cdot s_{\hat{a}})$
13. Intervalul de încredere pentru parametrul α , estimat pentru un eșantion observat, este definit de relația:
- a. $(b \pm t_{\alpha/2} \cdot s_{\hat{\beta}})$
 - b. $(a \pm t_{\alpha/2} \cdot s_{\hat{a}})$
 - c. $(b \pm t_{\alpha/2} \cdot s_{\hat{a}})$
14. Testarea semnificației parametrului β presupune formularea următoarelor ipoteze statistice
- a. $H_0 : b = 0; H_1 : b \neq 0$
 - b. $H_0 : \beta = 0; H_1 : \beta \neq 0$
 - c. $H_0 : \beta \neq 0; H_1 : \beta = 0$
15. Pentru testarea semnificației parametrului β se folosește
- a. statistica t Student
 - b. statistica F

- c. statistica χ^2
16. Coeficientul de corelație Pearson se folosește în cazul unei regresii
 - a. liniare simple
 - b. liniare multiple
 - c. neliniare, simple sau multiple
17. Coeficientul de corelație Pearson poate lua valori în intervalul:
 - a. $[-3, +3]$
 - b. $[0, 1]$
 - c. $[-1, +1]$
18. Pentru testarea semnificației coeficientului de corelație Pearson se folosește
 - a. statistica t Student
 - b. statistica F
 - c. statistica χ^2
19. Raportul de corelație se folosește în cazul unei regresii
 - a. liniare simple
 - b. liniare multiple
 - c. neliniare, simple sau multiple
20. Raportul de corelație poate lua valori în intervalul
 - a. $[-3, +3]$
 - b. $[0, 1]$
 - c. $[-1, +1]$
21. Pentru testarea semnificației raportului de corelație Pearson se folosește
 - a. statistica t Student
 - b. statistica F
 - c. statistica χ^2
22. Raportul de determinație arată
 - a. proporția variației variabilei dependente explicate prin modelul de regresie
 - b. proporția variației variabilei independente explicate prin modelul de regresie
 - c. proporția variației variabilei reziduale explicate prin modelul de regresie
23. Raportul de determinație poate lua valori în intervalul
 - a. $[-3, +3]$
 - b. $[0, 1]$
 - c. $[-1, +1]$

24. Coeficientul de concordanță Fechner se calculează după relația

a. $K = \frac{C - D}{C + D}$

b. $\theta = 1 - \frac{6 \cdot \sum d_i^2}{n \cdot (n^2 - 1)}$

c. $\tau = \frac{2 \cdot S}{n \cdot (n - 1)}$

d. $Q = \frac{ad - bc}{ad + bc}$

25. Coeficientul de concordanță Fechner poate lua valori în intervalul

a. $[-3, +3]$

b. $[0, 1]$

c. $[-1, +1]$

26. Coeficientul de corelație a rangurilor Spearman se calculează după relația

a. $K = \frac{C - D}{C + D}$

b. $\theta = 1 - \frac{6 \cdot \sum d_i^2}{n \cdot (n^2 - 1)}$

c. $\tau = \frac{2 \cdot S}{n \cdot (n - 1)}$

d. $Q = \frac{ad - bc}{ad + bc}$

27. Coeficientul de corelație a rangurilor Spearman poate lua valori în intervalul

a. $[-3, +3]$

b. $[0, 1]$

c. $[-1, +1]$

28. Coeficientul de corelație a rangurilor Kendall se calculează după relația

a. $K = \frac{C - D}{C + D}$

b. $\theta = 1 - \frac{6 \cdot \sum d_i^2}{n \cdot (n^2 - 1)}$

$$c. \tau = \frac{2 \cdot S}{n \cdot (n-1)}$$

$$d. Q = \frac{ad - bc}{ad + bc}$$

29. Coeficientul de corelație a rangurilor Kendall poate lua valori în intervalul

a. $[-3, +3]$

b. $[0, 1]$

c. $[-1, +1]$

30. Coeficientul de asociere caracterizează legătura dintre

a. două variabile numerice

b. două variabile nominale

c. o variabilă numerică și o variabilă nominală

31. Coeficientul de asociere Yule se calculează după relația

$$a. K = \frac{C - D}{C + D}$$

$$b. \theta = 1 - \frac{6 \cdot \sum d_i^2}{n \cdot (n^2 - 1)}$$

$$c. \tau = \frac{2 \cdot S}{n \cdot (n-1)}$$

$$d. Q = \frac{ad - bc}{ad + bc}$$

32. Coeficientul de asociere Yule poate lua valori în intervalul

a. $[-3, +3]$

b. $[0, 1]$

c. $[-1, +1]$



13.8. Recunoașteți exemplul corect

1. În urma observării și prelucrării datelor la nivelul unui eșantion de volum $n=50$ se obține o ecuație de regresie de forma: $y_x=15-1,5x$. Varianța variabilei Y la nivelul populației este egală cu 25. Pentru testarea coeficientului de regresie β , se folosește
 - a. testul t
 - b. testul F
 - ☒ c. testul z
2. Pentru un eșantion de volum $n=10$ se obține o ecuație de regresie de forma: $y_x=15+1,5x$. Varianța variabilei Y la nivelul populației este egală cu 25. Pentru testarea coeficientului de regresie β , se folosește
 - ☒ a. testul t
 - b. testul z
 - c. testul F
3. La nivelul unui eșantion se obține o valoare calculată a statisticii test $t=5,821$. Dacă valoarea teoretică a statisticii t Student este $t_{0,025,3}=3.182$, atunci se poate considera că
 - a. valoarea parametrului de regresie β este semnificativ diferită de zero
 - b. valoarea parametrului de regresie β este egală cu zero
 - ☒ c. există o legătură statistică între variabilele corelate
4. Pe baza unui eșantion de volum $n=100$, s-a estimat un model de regresie de forma: $y_x=6-1,5x$. În acest caz, se poate considera că
 - a. nivelul variabilei Y crește, în medie, cu 1,5 unități la o creștere cu o unitate a nivelului variabilei X
 - ☒ b. nivelul variabilei Y scade, în medie, cu 1,5 unități la o creștere cu o unitate a nivelului variabilei X
 - c. nivelul variabilei X scade, în medie, cu 1,5 unități la o creștere cu o unitate a nivelului variabilei Y
5. Pentru un model de regresie de forma: $y_x=6+1,5x$, valoarea coeficientului de regresie b arată că
 - ☒ a. nivelul variabilei Y crește, în medie, cu 1,5 unități la o creștere cu o unitate a nivelului variabilei X
 - b. nivelul variabilei Y scade, în medie, cu 1,5 unități la o creștere cu o unitate a nivelului variabilei X

- c. nivelul variabilei X crește, în medie, cu 1,5 unități la o creștere cu o unitate a nivelului variabilei Y
6. Pentru testarea semnificației coeficientului de corelație Pearson s-a obținut la nivelul unui eșantion o valoare calculată a statisticii test $t=9,52$. Dacă valoarea teoretică a statisticii test t Student este $t_{0,025,48}=1.96$, atunci se poate considera că
- valoarea parametrului coeficientul de corelație ρ este semnificativ diferită de zero
 - valoarea parametrului coeficientul de corelație ρ este egală cu zero
 - există o legătură statistică între variabilele corelate
7. Pentru o valoare estimată a coeficientului de corelație egală cu $r_{y_x} = -0,953$, se poate considera că
- între variabilele analizate există o legătură directă puternică
 - între variabilele analizate există o legătură inversă slabă
 - între variabilele analizate există o legătură inversă puternică
8. O valoare estimată a coeficientului de corelație este egală cu $r_{y_x} = +0,953$. În acest caz, se poate considera că
- între variabilele analizate există o legătură directă puternică
 - între variabilele analizate există o legătură directă slabă
 - între variabilele analizate există o legătură inversă puternică
9. Pentru un eșantion de volum $n=15$, s-a estimat un model de regresie de forma: $y_x = a + bx$ și valoarea raportului de corelație. Pentru testarea raportului de corelație, s-a obținut o valoare calculată a statisticii test $F=7,52$. Pentru un risc de 0,05, se poate considera că ~~Testarea~~ $F_{0,05,13,1} = 4,6$
- raportul de corelație este semnificativ diferit de zero
 - raportul de corelație este egal cu zero
 - raportul de corelație nu este semnificativ diferit de zero
10. Pentru un eșantion de volum $n=15$, s-a estimat un model de regresie de forma: $y_x = a + bx$ și valoarea raportului de corelație. Pentru testarea raportului de corelație, s-a obținut o valoare calculată a statisticii test $F=2,52$. Pentru un risc de 0,05, se poate considera că
- raportul de corelație este semnificativ diferit de zero
 - raportul de corelație este egal cu zero
 - raportul de corelație nu este semnificativ diferit de zero
11. Pentru un model de regresie simplă liniară se obține o valoare estimată a raportului de determinație egală cu $d_{y_x} = 0,832$. În acest caz, se poate considera că
- 83,2% din variația variabilei independente X poate fi explicată prin variația variabilei dependente Y

- b. 83,2% din variația variabilei dependente Y poate fi explicată prin variația variabilei aleatoare
- c. 83,2% din variația variabilei dependente Y poate fi explicată prin variația variabilei independente X
12. Considerând ecuația de regresie $y_x = 2 + 5x$, valoarea estimată a variabilei dependente Y pentru o valoare $x = 7$ este
- a. 7
- b. 17
- c. 37
13. În urma unui studiu realizat asupra distribuției a două variabile corelate, X și Y, s-au obținut următoarele rezultate în SPSS:

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,588	,348		1,687	,100
	VARSTA	,198	,010	,958	20,564	,000

a. Dependent Variable: SALARIU

Pe baza datelor din output, se poate considera că modelul de regresie estimat este de forma:

a. $y = 0,198 + 0,588x$

b. $y = 0,588 + 0,198x$

c. $y = 0,348 + 0,010x$

14. În vederea testării semnificației coeficientului de regresie β , s-au înregistrat valorile a două variabile corelate, X și Y. În urma prelucrării datelor în SPSS s-au obținut următoarele rezultate

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,588	,348		1,687	,100
	VARSTA	,198	,010	,958	20,564	,000

a. Dependent Variable: SALARIU

Pe baza datelor din output, se poate considera că

- valoarea parametrului de regresie β este semnificativ diferită de zero, în condițiile unui risc de 5%
- valoarea parametrului de regresie β este egală cu zero, în condițiile unui risc de 5%
- valoarea parametrului de regresie β este semnificativ diferită de zero, în condițiile unui risc de 95%

15. În urma unui studiu realizat asupra distribuției a două variabile corelate, X și Y , s-au obținut următoarele rezultate în SPSS:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,958 ^a	,918	,915	,6351

a. Predictors: (Constant), VARSTA

Pe baza datelor din output, se poate considera că

- între variabilele analizate există o legătură puternică
- între variabilele analizate există o legătură slabă
- între variabilele analizate nu există legătură



13.9. Răspunsuri corecte

13.7. Alegeți răspunsul corect

1. a) Analiza de regresie este o metodă statistică care permite studierea și măsurarea relației care există între două sau mai multe variabile, precum și descoperirea legii relative la forma legăturilor dintre variabile.
2. c) Analiza de corelație este o metodă care permite măsurarea intensității legăturii dintre variabile.
3. a) Variabila factorială (independentă) este variabila care explică variația altei variabile.
4. b) Variabila rezultativă (dependentă) este variabila estimată, Y .
5. c) Variabila reziduală (aleatoare) este variabila care sintetizează toți factorii care influențează variabila Y , dar sunt nespecificați în model.
6. a) și b) După numărul variabilelor corelate, legăturile statistice pot fi legături simple și legături multiple.
7. c) și d) După sensul legăturii statistice, acestea pot fi legături directe și legături inverse.
8. b) și c) După forma legăturii statistice, acestea pot fi legături rectilinii și legături curbilinii.
9. a), b), c) și d) Metodele elementare de analiză a legăturilor statistice sunt: metoda seriilor statistice paralele interdependente, metoda grupărilor statistice, metoda tabelului de corelație și metoda grafică.
10. a) Metoda celor mai mici pătrate minimizează suma pătratelor erorilor.
11. a) Intervalul de încredere pentru parametrul β este probabilist.
12. a) Intervalul de încredere pentru parametrul β , estimat pentru un eșantion observat, este definit de relația: $(b \pm t_{\alpha/2} \cdot s_{\hat{\beta}})$.
13. b) Intervalul de încredere pentru parametrul α , estimat pentru un eșantion observat, este definit de relația: $(a \pm t_{\alpha/2} \cdot s_{\hat{\alpha}})$.
14. b) Testarea semnificației parametrului β presupune formularea următoarelor ipoteze statistice: $H_0: \beta = 0$; $H_1: \beta \neq 0$.
15. a) Pentru testarea semnificației parametrului β se folosește statistica t Student.
16. a) Coeficientul de corelație Pearson se folosește în cazul unei regresii liniare simple.

17. c) Coeficientul de corelație Pearson poate lua valori în intervalul $[-1, +1]$.

18. a) Pentru testarea semnificației coeficientului de corelație Pearson se folosește statistica t Student.

19. a), b) și c) Raportul de corelație se folosește în cazul unei regresii liniare cât și neliniare, simple sau multiple.

20. b) Raportul de corelație poate lua valori în intervalul $[0, 1]$.

21. b) Pentru testarea semnificației raportului de corelație Pearson se folosește statistica F .

22. a) Raportul de determinație arată proporția variației variabilei dependente explicate prin modelul de regresie.

23. b) Raportul de determinație poate lua valori în intervalul $[0, 1]$.

24. a) Coeficientul de concordanță Fechner se calculează după relația:

$$K = \frac{C - D}{C + D}.$$

25. c) Coeficientul de concordanță Fechner poate lua valori în intervalul $[-1, +1]$.

26. b) Coeficientul de corelație a rangurilor Spearman se calculează după

$$6 \cdot \sum d_i^2$$

relația: $\theta = 1 - \frac{6 \cdot \sum d_i^2}{n \cdot (n^2 - 1)}$.

27. c) Coeficientul de corelație a rangurilor Spearman poate lua valori în intervalul $[-1, +1]$.

28. c) Coeficientul de corelație a rangurilor Kendall se calculează după relația:

$$\tau = \frac{2 \cdot S}{n \cdot (n - 1)}$$

29. c) Coeficientul de corelație a rangurilor Kendall poate lua valori în intervalul $[-1, +1]$.

30. b) Coeficientul de asociere caracterizează legătura dintre două variabile nominale.

31. d) Coeficientul de asociere Yule se calculează după relația: $Q = \frac{ad - bc}{ad + bc}$.

32. c) Coeficientul de asociere Yule poate lua valori în intervalul $[-1, +1]$.

13.8. Recunoașteți exemplul corect

1. c) În urma observării și prelucrării datelor la nivelul unui eșantion de volum $n=50$ se obține o ecuație de regresie de forma: $y_x=15-1,5x$. Varianța variabilei Y la nivelul populației este egală cu 25. Pentru testarea coeficientului de regresie β , se folosește testul z .
2. a) Pentru un eșantion de volum $n=10$ se obține o ecuație de regresie de forma: $y_x=15+1,5x$. Varianța variabilei Y la nivelul populației este egală cu 25. Pentru testarea coeficientului de regresie β , se folosește testul t .
3. a) și c) La nivelul unui eșantion se obține o valoare calculată a statisticii test $t=5,821$. Dacă valoarea teoretică a statisticii t Student este $t_{0,025,3}=3,182$, atunci se poate considera că valoarea parametrului de regresie β este semnificativ diferită de zero, deci există o legătură statistică între variabilele corelate.
4. b) Pe baza unui eșantion de volum $n=100$, s-a estimat un model de regresie de forma: $y_x=6-1,5x$. În acest caz, se poate considera că nivelul variabilei Y scade, în medie, cu 1,5 unități la o creștere cu o unitate a nivelului variabilei X .
5. a) Pentru un model de regresie de forma: $y_x=6+1,5x$, valoarea coeficientului de regresie b arată că nivelul variabilei Y crește în medie cu 1,5 unități la o creștere cu o unitate a nivelului variabilei X .
6. a) și c) Pentru testarea semnificației coeficientului de corelație Pearson s-a obținut la nivelul unui eșantion o valoare calculată a statisticii test $t=9,52$. Dacă valoarea teoretică a statisticii t Student este $t_{0,025,48}=1,96$, atunci se poate considera că valoarea parametrului coeficientul de corelație ρ este semnificativ diferită de zero, deci există o legătură statistică între variabilele corelate.
7. c) Pentru o valoare estimată a coeficientului de corelație egală cu $r_{yx} = -0,953$, se poate considera că între variabilele analizate există o legătură inversă puternică.
8. a) O valoare estimată a coeficientului de corelație este egală cu $r_{yx} = +0,953$. În acest caz, se poate considera că între variabilele analizate există o legătură directă puternică.
9. a) Pentru un eșantion de volum $n=15$, s-a estimat un model de regresie de forma: $y_x=a+bx$ și valoarea raportului de corelație. Pentru testarea raportului de corelație, s-a obținut o valoare calculată a statisticii test $F=7,52$. Valoarea teoretică a statisticii test F se citește din Tabela Fisher pentru un risc $\alpha=0,05$ și $v_1=k-1=2-1=1$, $v_2=n-k=15-2=13$ grade de libertate, $F_{0,05,1,13}=4,667$. Deoarece $(F_{\text{calc}}=7,52) > (F_{0,05,1,13}=4,667)$, se acceptă ipoteza prin care se admite că raportul de corelație este semnificativ diferit de zero.
10. c) și b) Pentru un eșantion de volum $n=15$, s-a estimat un model de regresie de forma: $y_x=a+bx$ și valoarea raportului de corelație. Pentru testarea raportului

de corelație, s-a obținut o valoare calculată a statisticii test $F=2,52$. Valoarea teoretică a statisticii test F se citește din Tabela Fisher pentru un risc $\alpha=0,05$ și $v_1=k-1=2-1=1$, $v_2=n-k=15-2=13$ grade de libertate, $F_{0,05,1,13}=4,667$. Deoarece $F_{\text{calc}}=2,52 < F_{0,05,1,13}=4,667$, se acceptă ipoteza prin care se admite că raportul de corelație nu este semnificativ diferit de zero.

11. c) Pentru un model de regresie simplă liniară se obține o valoare estimată a raportului de determinație egală cu $d_{y_x}=0,832$. În acest caz, se poate considera că 83,2% din variația variabilei dependente Y poate fi explicată prin variația variabilei independente X .

12. c) Considerând ecuația de regresie, $y_x=2+5x$, valoarea estimată a variabilei dependente Y pentru o valoare $x=7$, este: $y_x=2+5*7=37$.

13. b) În urma unui studiu realizat asupra distribuției a două variabile corelate, X și Y , s-au obținut următoarele rezultate în SPSS:

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,588	,348		1,687	,100
	VARSTA	,198	,010	,958	20,564	,000

a. Dependent Variable: SALARIU

Pe baza datelor din output, se poate considera că modelul de regresie estimat este de forma: $y=0,588+0,198x$.

14. În vederea testării semnificației coeficientului de regresie β , s-au înregistrat valorile a două variabile corelate, X și Y . În urma prelucrării datelor în SPSS s-au obținut următoarele rezultate:

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,588	,348		1,687	,100
	VARSTA	,198	,010	,958	20,564	,000

a. Dependent Variable: SALARIU

Valoarea $\text{Sig.} = 0,000 < 0,05$, se respinge ipoteza H_0 ($H_0: \beta = 0$). Pe baza datelor din output, se poate considera că valoarea parametrului de regresie β este semnificativ diferită de zero, în condițiile unui risc de 5%.

15. a) În urma unui studiu realizat asupra distribuției a două variabile corelate, X și Y , s-au obținut următoarele rezultate în SPSS:

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,958 ^a	,918	,915	,6351

a. Predictors: (Constant), VARSTA

Pe baza datelor din output, se poate considera că între variabilele analizate există o legătură puternică ($\eta^2 = 0,918$).

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,958 ^a	,918	,915	,6351

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,958 ^a	,918	,915	,6351



13.10. Probleme

13.10.1. Estimarea parametrilor modelului de regresie

1. În urma unui studiu realizat asupra distribuției a două variabile corelate, X și Y , s-au obținut următoarele rezultate: $\sigma_x=4$, $\sigma_y=0,6$, $r_{yx} = -0,872$. Se cere să se estimeze valoarea coeficientului de regresie β și să se interpreteze rezultatul obținut.

Între coeficientul de regresie β și coeficientul de corelație r_{yx} există relația:

$$r_{yx} = b \cdot \frac{\sigma_x}{\sigma_y}, \text{ de unde coeficientul de}$$

regresie β se poate afla după relația:

$$\beta = r_{yx} \cdot \frac{\sigma_y}{\sigma_x}$$

$$y = a + bx$$

Coeficientul de regresie β este:

$$\beta = (-0,872) \cdot \frac{0,6}{4} = -0,1308$$

Interpretare: Semnul parametrului β arată că între cele două variabile corelate, X și Y , există o legătură inversă. Valoarea parametrului β arată că la o creștere cu o unitate a lui X , Y scade în medie cu 0,1308 unități.

2. Pentru un eșantion de volum $n=102$ se cunosc următoarele rezultate: $\sum e_i^2=16$, $\sum (x_i - \bar{x})^2 = 25$, $y_x = 4 + 2,5x$. Se cere să se estimeze prin interval de încredere coeficientul de regresie β , considerând un risc $\alpha=0,05$.

Intervalul de încredere pentru parametrul β , în cazul unui eșantion de volum mare ($n>30$), este definit de relația:

$$(b \pm z_{\alpha/2} \cdot s_{\hat{\beta}}),$$

unde:

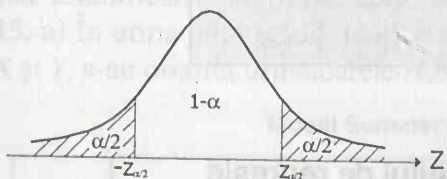
$z_{\alpha/2}$ este valoarea statisticii $Z \sim N(0,1)$ pentru un risc $\alpha/2$.

Estimația varianței erorilor este:

$$s_e^2 = \frac{16}{102-2} = 0,16$$

Estimația abaterii medii pătratice a estimatorului parametrului β este:

$$s_{\hat{\beta}} = \sqrt{\frac{0,16}{25}} = 0,08$$



Valoarea statisticii Z, corespunzătoare unui risc ales, se citește din tabelul Gauss Laplace.

$$s_{\hat{\beta}} = \sqrt{\frac{s_e^2}{\sum_i (x_i - \bar{x})^2}}$$

$$s_e^2 = \frac{\sum_i e_i^2}{n-2}$$

3. La nivelul unui eșantion de volum $n=30$ firme se cunosc următoarele rezultate: $y_x = 4 + 2x$, și $s_{\hat{\beta}}^2 = 0,64$. Se cere să se estimeze prin interval de încredere coeficientul de regresie β , considerând un risc $\alpha=0,10$.

Intervalul de încredere pentru parametrul β , în cazul unui eșantion de volum mic ($n \leq 30$), este:

$$(b \pm t_{\alpha/2, n-2} \cdot s_{\hat{\beta}}),$$

unde:

$t_{\alpha/2, n-2}$ este valoarea statisticii t care urmează o distribuție de probabilitate Student cu $(n-2)$ grade de libertate, $t \sim t(n-2)$.

$$s_{\hat{\beta}} = \sqrt{s_{\hat{\beta}}^2}$$

Valoarea teoretică a statisticii Z:

Valoarea teoretică a statisticii Z se citește din tabela Gauss-Laplace pentru un risc $\alpha=0,05$. Aceasta este $z_{\alpha/2}=1,96$.

Intervalul de încredere pentru parametrul β este:

$$(2,5 \pm 1,96 \cdot 0,08) = (2,34; 2,66)$$

Interpretare:

Cu o probabilitate de 0,95 se poate garanta că valoarea parametrului β este acoperită de intervalul 2,34 și 2,66.

Estimația abaterii medii pătratice a estimatorului parametrului β este:

$$s_{\hat{\beta}} = \sqrt{0,64} = 0,8$$

Valoarea teoretică a statisticii t Student:

Valoarea teoretică a statisticii t Student se citește din tabela Student pentru un risc $\alpha=0,10$ și $v=n-2=28$ grade de libertate: $t_{0,05;28}=1,701$.

α v	...	$t_{0.05}$	...
...	...	...	...
28	...	1,701	...
...	...	...	...

Intervalul de încredere pentru parametrul β este:

$$(2 \pm 1,701 \cdot 0,8) = (0,64; 3,36)$$

Interpretare: Cu o probabilitate de 0,90 se poate garanta că valoarea parametrului β este cuprinsă între 0,64 și 3,36.

4. Pe baza datelor observate pe un eșantion de volum $n=18$, se cunosc următoarele rezultate: $\sum e_i^2=16$; $\sum (x_i - \bar{x})^2 = 25$; $y_x=4-2,5x$. Se cere să se estimeze prin interval de încredere coeficientul de regresie β , considerând un risc $\alpha=0,05$.

Intervalul de încredere pentru parametrul β , în cazul unui eșantion de volum mic ($n \leq 30$), este:

$$(b \pm t_{\alpha/2, n-2} \cdot s_{\beta}),$$

unde:

$t_{\alpha/2, n-2}$ este valoarea statisticii t care urmează o distribuție de probabilitate Student cu $(n-2)$ grade de libertate, $t \sim t(n-2)$.

$$s_{\beta} = \sqrt{\frac{s_e^2}{\sum_i (x_i - \bar{x})^2}},$$

Estimația varianței erorilor este:

$$s_e^2 = \frac{16}{18-2} = 1$$

Estimația abaterii medii pătratice a estimatorului parametrului β este:

$$s_{\beta} = \sqrt{\frac{1}{25}} = 0,2$$

Valoarea teoretică a statisticii t Student:

Valoarea teoretică a statisticii t Student se citește din tabela Student pentru un risc $\alpha=0,05$ și $v=n-2=18-2=16$ grade de libertate: $t_{0,025;16}=2,12$.

$$s_e^2 = \frac{\sum_i e_i^2}{n-2}.$$

α v	...	$t_{0.025}$	...
16	...	2,12	...

Intervalul de încredere pentru parametrul β este:

$$(-2,5 \pm 2,12 \cdot 0,2) = (-2,92; -2,07).$$

Interpretare: Cu o probabilitate de 0,95 se poate garanta că valoarea parametrului β este acoperită de intervalul: $(-2,92; -2,07)$.

5. Pentru două variabile corelate, observate pe un eșantion de volum $n=30$, se cunosc următoarele rezultate: $y_x = a + bx$, $\sum_i y_i = 300$, $\sum_i x_i = 180$, $b=0,5$. Se cere să se estimeze nivelul variabilei Y pentru o valoare $x=10$.

Modelul de regresie este:

$$y_x = a + b \cdot x_i,$$

unde:

$$a = \bar{y} - b \cdot \bar{x}.$$

Valoarea estimată a variabilei Y pentru $x=10$ este:

$$y_{x=10} = a + 10b.$$

Elemente de calcul pentru coeficientul de regresie a :

$$\bar{y} = \frac{300}{30} = 10, \bar{x} = \frac{180}{30} = 6$$

$$a = 10 - 0,5 \cdot 6 = 7$$

Estimarea nivelului variabilei Y pentru o valoare a lui $x=10$ este:

$$y_x = 7 + 0,5 \cdot 10 = 12$$

Interpretare:

Legătura dintre cele două variabile este o legătură directă ($\beta > 0$). La o valoare a lui X egală cu 10, Y va atinge o valoare egală cu 12.

13.10.2. Testarea semnificației coeficientului de regresie β

1. La nivelul unui eșantion de volum $n=18$, se cunosc următoarele rezultate: $y_x=4-2,5x$; $\sum e_i^2=16$, $\sum (x_i - \bar{x})^2 = 25$. Se cere să se testeze dacă valoarea coeficientului de regresie β este semnificativ diferită de zero, considerând un risc $\alpha=0,05$.

Ipoteze statistice:

$$H_0: \beta = 0$$

$$H_1: \beta \neq 0$$

Statistica test, când $n < 30$, se calculează după relația:

$$t = \frac{b}{s_{\beta}},$$

unde:

$$s_{\beta} = \sqrt{\frac{s_e^2}{\sum (x_i - \bar{x})^2}},$$

$$s_e^2 = \frac{\sum e_i^2}{n-2}.$$

Regula de decizie:

Dacă $|t_{calc}| > t_{\alpha/2, n-2} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $|t_{calc}| \leq t_{\alpha/2, n-2} \Rightarrow$ Se acceptă ipoteza H_0 .

Estimația varianței erorilor este:

$$s_e^2 = \frac{16}{18-2} = 1$$

Estimația abaterii medii pătratice a estimatorului parametrului β este:

$$s_{\beta} = \sqrt{\frac{1}{25}} = 0,2$$

Statistica test calculată:

$$t_{calc} = \frac{-2,5}{0,2} = -12,5$$

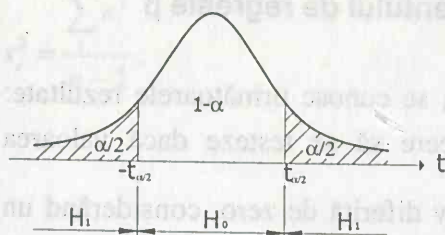
Valoarea teoretică a statisticii test:

Valoarea teoretică a statisticii t Student se citește din tabela Student pentru un risc $\alpha=0,05$ și $v=n-2=18-2=16$ grade de libertate: $t_{0,025;16}=2,12$.

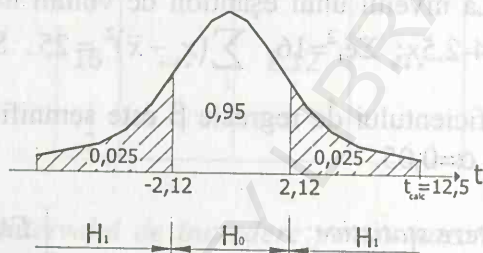
α v	...	$t_{0.025}$	...
...	...	...	...
16	...	2,12	...
...	...	...	...

Decizie:

Deoarece $|t_{calc}| = 12,5 > t_{teor} = 2,12$, se



respinge ipoteza H_0 , se acceptă ipoteza H_1 .



Interpretare:

Parametrul β este semnificativ diferit de zero, între variabilele X și Y există o legătură inversă. Se poate garanta acest rezultat cu o probabilitate de 0,95.

2. În cazul a două variabile corelate, observate pe un eșantion de volum $n=50$, se cunosc următoarele rezultate: $y_x=4+2,5x$; $\sum e_i^2=16$; $\sum (x_i - \bar{x})^2 = 25$. Se cere să se testeze dacă valoarea coeficientului de regresie β este semnificativ diferită de zero, considerând un risc $\alpha=0,05$.

Ipoteze statistice:

$H_0: \beta = 0$

$H_1: \beta \neq 0$

Statistica test, când $n>30$, se calculează după relația:

$$z = \frac{b}{s_{\hat{\beta}}},$$

unde:

$$s_{\hat{\beta}} = \sqrt{\frac{s_e^2}{\sum_i (x_i - \bar{x})^2}},$$

$$s_e^2 = \frac{\sum_i e_i^2}{n-2}.$$

Estimația varianței erorilor este:

$$s_e^2 = \frac{16}{25-2} = 0,696$$

Estimația abaterii medii pătratice a estimatorului parametrului β este:

$$s_{\hat{\beta}} = \sqrt{\frac{0,696}{25}} = 0,167$$

Statistica test calculată:

$$z_{calc} = \frac{2,5}{0,167} = 14,97$$

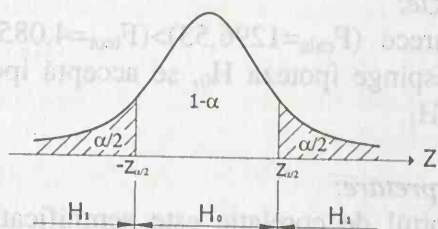
Valoarea teoretică a statisticii test:

Valoarea teoretică a statisticii test Z se

Regula de decizie:

Dacă $|z_{calc}| > z_{teor} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $|z_{calc}| \leq z_{teor} \Rightarrow$ Se acceptă ipoteza H_0 .



citește din tabelul Gauss-Laplace pentru un risc $\alpha=0,05$. Aceasta este $z_{teor}=1,96$.

Decizie:

Deoarece $|z_{calc}| = 14,97 > z_{teor} = 1,96$, se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Interpretare:

Parametrul β este semnificativ diferit de zero, între variabilele X și Y există o legătură directă. Se poate garanta acest rezultat cu o probabilitate de 0,95.

13.10.3. Analiza de corelație

1. În urma prelucrării datelor observate la nivelul unui eșantion format din 50 unități a rezultat un model de regresie linear $y_x=a+bx$. Valoarea estimată a raportului de corelație este 0,982. Se cere să se testeze semnificația raportului de corelație la nivelul populației, asumându-ne un risc de 0,05.

Ipoteze statistice:

$H_0: \eta = 0$

$H_1: \eta > 0$

Statistica test calculată:

$$F = \frac{n-k}{k-1} \cdot \frac{\eta_{y_x}^2}{1-\eta_{y_x}^2}$$

Regula de decizie:

Dacă $F_{calc} > F_{teor} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Valoarea calculată a statisticii test F este:

$$F_{calc} = \frac{50-2}{2-1} \cdot \frac{0,982^2}{1-0,982^2} = 1296,53$$

Valoarea teoretică a statisticii test:

Valoarea teoretică a statisticii test F se citește din Tabela Fisher pentru un risc $\alpha=0,05$ și $v_1=k-1=1$, $v_2=n-k=48$ grade de libertate: $F_{teor}=4,085$.

Dacă $F_{calc} \leq F_{teor} \Rightarrow$ Se acceptă ipoteza H_0 .

$V_1 \backslash V_2$	1	...	...
...			
40	4,085	...	...
...			

Decizie:

Deoarece $(F_{calc}=1296,53) > (F_{teor}=4,085)$, se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Interpretare:

Raportul de corelație este semnificativ diferit de zero, între variabilele X și Y există o legătură puternică. Se poate garanta acest rezultat cu o probabilitate de 0,95.

2. La nivelul unui eșantion de volum $n=102$ firme, pentru două variabile corelate liniar, s-a obținut o valoare estimată a raportului de corelație egală cu 0,90. Considerând un risc $\alpha=0,10$, se cere să se testeze semnificația raportului de corelație.

Ipoteze statistice:

$H_0: \eta = 0$

$H_1: \eta > 0$

Statistica test calculată:

$$F = \frac{n-k}{k-1} \cdot \frac{\eta_{y_x}^2}{1-\eta_{y_x}^2}$$

Regula de decizie:

Dacă $F_{calc} > F_{teor} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $F_{calc} \leq F_{teor} \Rightarrow$ Se acceptă ipoteza H_0 .

Valoarea calculată a statisticii test F este:

$$F_{calc} = \frac{102-2}{2-1} \cdot \frac{0,90^2}{1-0,90^2} = 426,32$$

Valoarea teoretică a statisticii test:

Valoarea teoretică a statisticii test F se citește din Tabela Fisher pentru un risc $\alpha=0,10$ și $v_1=k-1=1$, $v_2=n-k=100$ grade de libertate: $F_{teor}=2,748$.

Regula de decizie:

Dacă $|t_{calc}| > t_{\alpha/2, n-2} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $|t_{calc}| \leq t_{\alpha/2, n-2} \Rightarrow$ Se acceptă ipoteza H_0 .

α v	...	$t_{0.025}$	...
...			
28	...	2,048	...
...			

Decizie:

Deoarece $(t_{calc}=32,52) > (t_{0,025,28}=2,048)$, se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Interpretare:

Coeficientul de corelație este semnificativ diferit de zero, între variabilele X și Y există o legătură directă puternică. Se poate garanta acest rezultat cu o probabilitate de 0,95.

4. La nivelul unui eșantion de volum $n=81$ firme, pentru două variabile corelate liniar, s-a obținut un coeficient de corelație $r_{yx} = -0,90$. Considerând un risc $\alpha=0,05$, se cere să se testeze semnificația coeficientului de corelație.

Ipoteze statistice:

$H_0: \rho = 0$

$H_1: \rho \neq 0$

Statistica test calculată:

$$z = \frac{r_{yx}}{s_{\rho}} = \frac{r_{yx} \cdot \sqrt{n-2}}{\sqrt{1-r_{yx}^2}}$$

Regula de decizie:

Dacă $|z_{calc}| > z_{\alpha/2} \Rightarrow$ Se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Dacă $|z_{calc}| \leq z_{\alpha/2} \Rightarrow$ Se acceptă ipoteza H_0 .

Valoarea calculată a statisticii test Z este:

$$z_{calc} = \frac{-0,90 \cdot \sqrt{81-2}}{\sqrt{1-0,90^2}} = -18,35$$

Valoarea teoretică a statisticii test:

Valoarea teoretică a statisticii test z se citește din tabelul Gauss-Laplace pentru un risc $\alpha=0,05$. Aceasta este $z_{teor}=1,96$.

Decizie:

Deoarece $|z_{calc}| = 18,35 > z_{teor} = 1,96$, se respinge ipoteza H_0 , se acceptă ipoteza H_1 .

Interpretare:

Coeficientul de corelație este semnificativ diferit de zero, între variabilele X și Y există o legătură inversă puternică. Se garantează acest rezultat cu o probabilitate de 0,95.

13.10.4. Corelația neparametrică

1. Pentru rangurile a două variabile corelate, X și Y , înregistrate la nivelul unui eșantion format din 15 unități, s-a obținut pentru suma pătratelor diferențelor dintre ranguri următorul rezultat: $\sum d_i^2 = 25$. Să se estimeze coeficientul de corelație neparametric Spearman.

Coeficientul de corelație neparametric Spearman se calculează după relația:

$$\theta = 1 - \frac{6 \cdot \sum d_i^2}{n \cdot (n^2 - 1)},$$

unde:

$d_i = R_{x_i} - R_{y_i}$, unde

R_{x_i} este rangul valorilor variabilei X ;

R_{y_i} este rangul valorilor variabilei Y ;

n este numărul perechilor $(x_i; y_i)$.

Valoarea coeficientului de corelație Spearman este:

$$\theta = 1 - \frac{6 \cdot 25}{15 \cdot (15^2 - 1)} = 1 - 0,045 = 0,955$$

Interpretare: Semnul coeficientului de corelație Spearman arată că între variabilele X și Y există o legătură directă ($\theta > 0$). Valoarea coeficientului evidențiază o legătură puternică.

2. Pentru un eșantion de 7 firme s-a observat dinamica cifrei de afaceri (Y) și dinamica numărului de angajați (X), obținându-se următoarele rezultate:

Firma	X	Y
A	13	12
B	22	15
C	28	24
D	26	35
E	20	24
F	25	40
G	10	10
Total	-	-

Se cere să se măsoare intensitatea legăturii dintre cele două variabile, folosind coeficientul de corelație a rangurilor Spearman.

Coeficientul de corelație a rangurilor Spearman se calculează după relația:

$$\theta = 1 - \frac{6 \cdot \sum d_i^2}{n \cdot (n^2 - 1)},$$

unde:

$$d_i = R_{x_i} - R_{y_i}, \text{ unde}$$

R_{x_i} este rangul valorilor variabilei

X;

R_{y_i} este rangul valorilor variabilei

Y;

n este numărul perechilor $(x_i; y_i)$.

Elementele de calcul sunt prezentate în tabelul 13.1.

Valoarea coeficientului de corelație Spearman este:

$$\theta = 1 - \frac{6 \cdot 14}{7 \cdot (7^2 - 1)} = 1 - 0,25 = 0,75$$

Interpretare:

Semnul coeficientului de corelație Spearman arată că între variabilele X și Y există o legătură directă ($\theta > 0$).

Valoarea coeficientului evidențiază o legătură puternică (valoare apropiată de +1).

Tabelul 13.1. Elemente de calcul

Firma	R_{x_i}	R_{y_i}	d_i	d_i^2
C	1	3	-2	4
D	2	2	0	0
F	3	1	2	4
B	4	4	0	0
E	5	3	2	4
A	6	5	1	1
G	7	6	1	1
TOTAL	-	-	-	14

3. Pentru un eșantion de 7 firme s-a observat dinamica cifrei de afaceri (Y) și dinamica numărului de angajați (X), obținându-se următoarele rezultate:

Firma	X	Y
A	40	13
B	30	22
C	35	18
D	18	26
E	22	20
F	24	25
G	45	10
Total	-	-

Se cere să se măsoare intensitatea legăturii dintre cele două variabile, folosind coeficientul de corelație a rangurilor Spearman.

Coeficientul de corelație a rangurilor Spearman se calculează după relația:

$$\theta = 1 - \frac{6 \cdot \sum d_i^2}{n \cdot (n^2 - 1)},$$

unde:

$$d_i = R_{x_i} - R_{y_i}, \text{ unde}$$

R_{x_i} este rangul valorilor variabilei X;

R_{y_i} este rangul valorilor variabilei Y;

n este numărul perechilor (xi; yi).

Elementele de calcul sunt prezentate în tabelul 13.2.

Valoarea coeficientului de corelație Spearman este:

$$\theta = 1 - \frac{6 \cdot 106}{7 \cdot (7^2 - 1)} = 1 - 1,89 = -0,89$$

Interpretare:

Semnul coeficientului de corelație Spearman arată că între variabilele X și Y există o legătură inversă ($\theta < 0$). Valoarea coeficientului evidențiază o legătură puternică.

Tabelul 13.2. Elemente de calcul

Firma	R_{x_i}	R_{y_i}	d_i	d_i^2
G	1	7	-6	36
A	2	6	-4	16
C	3	5	-2	4
B	4	3	1	1
F	5	2	3	9
E	6	4	2	4
D	7	1	6	36
Total	-	-	-	106

4. Pentru un eșantion de 7 firme s-a observat dinamica cifrei de afaceri (Y) și dinamica numărului de angajați (X), obținându-se următoarele rezultate:

Firma	X	Y
A	40	13
B	30	22
C	35	18
D	18	26
E	22	20
F	24	25
G	45	10
Total	-	-

Se cere să se măsoare intensitatea legăturii dintre cele două variabile, folosind coeficientul de corelație a rangurilor Kendall.

Coeficientul de corelație a rangurilor Kendall se calculează după relația:

$$\tau = \frac{2 \cdot S}{n \cdot (n-1)},$$

unde:

$S = Q + P$. Valorile Q și P se află considerând rangurile variabilei Y ordonate după rangurile variabilei X .

Elementele de calcul sunt prezentate în tabelul 13.3.

Valoarea coeficientului de corelație Kendall este:

$$\tau = \frac{2 \cdot (-16)}{7 \cdot (7-1)} = -0,762$$

Interpretare:

Semnul coeficientului de corelație Kendall arată că între variabilele X și Y

Q – numărul de ranguri mai mici decât rangul considerat (se ia cu semnul minus);

P – numărul de ranguri mai mari decât rangul considerat (se ia cu semnul plus).

există o legătură inversă ($\tau < 0$). Valoarea coeficientului evidențiază o legătură puternică.

Tabelul 13.3. Elemente de calcul

Firma	R_{x_i}	R_{y_i}	Q	P	S
G	1	7	-6	0	-6
A	2	6	-5	0	-5
C	3	5	-4	0	-4
B	4	3	-2	1	-1
F	5	2	-1	1	0
E	6	4	0	0	0
D	7	1	0	0	0
Total	-	-	-	-	-16

14.2. Ajustarea seriilor cronologice

Capitolul 14

ANALIZA ȘI EXTRAPOLAREA SERIILOR

CRONOLOGICE

- ◆ Componentele unei serii cronologice
- ◆ Ajustarea seriilor cronologice
- ◆ Ajustarea seriilor cronologice cu caracter sezonier
- ◆ Criterii de alegere a celui mai bun model de trend
- ◆ Prognoza statistică
- ◆ Alegeți răspunsul corect
- ◆ Recunoașteți exemplul corect
- ◆ Răspunsuri corecte
- ◆ Probleme

14.1. Componentele unei serii cronologice

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economică, București, 2002, pp. 422-426)

14.1.1. Componente ale unei serii cronologice

- Tendința
- Fluctuația ciclică
- Variația sezonieră
- Variația aleatoare

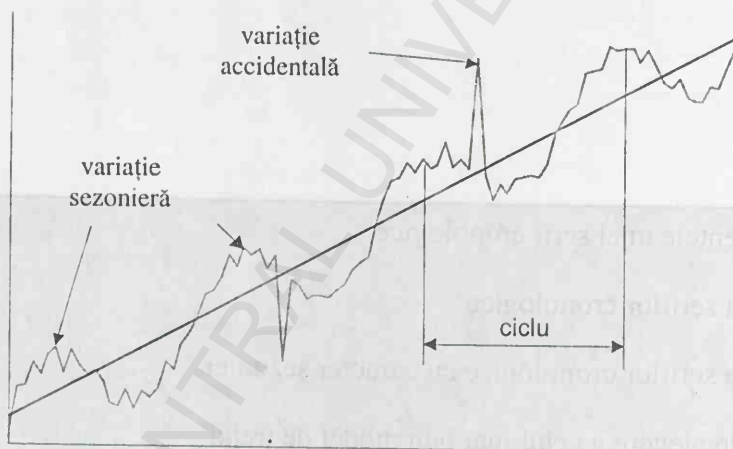


Figura 14.1. Componentele unei serii cronologice „complete”

14.1.2. Combinarea componentelor unei serii cronologice

- Modelul aditiv: $y_t = f_t + C_t + S_t + e_t$
- Modelul multiplicativ: $y_t = f_t \cdot C_t \cdot S_t \cdot e_t$
- Modelul mixt: $y_t = S_t + (f_t \cdot C_t \cdot e_t)$ sau $y_t = C_t + (f_t \cdot S_t \cdot e_t)$.

14.2. Ajustarea seriilor cronologice

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 426-444)

Ajustarea seriilor de timp presupune înlocuirea termenilor seriei date cu termeni teoretici care exprimă tendința de dezvoltare a unui fenomen.

• Metode elementare de ajustare

- **Metoda grafică.** Se aproximează vizual modelul de variație în timp care corespunde cel mai bine tendinței de dezvoltare a fenomenului. Aplicarea acestei metode constă în **construirea cronogramei și trasarea vizuală a liniei de tendință.**

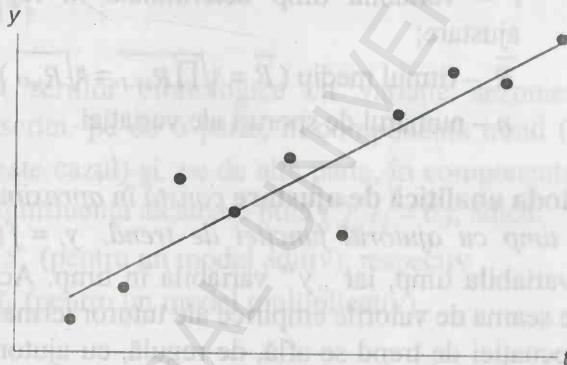


Fig.14.2. Ajustarea grafică

- **Metoda sporului mediu.** Ajustarea seriilor de timp pe baza sporului mediu se utilizează când termenii seriei variază aproximativ după o *progresie aritmetică*.

Termenii seriei ajustate se calculează după relația:

$$y_t = y_0 + t \cdot \bar{\Delta}$$

în care: y_t – reprezintă oricare termen ajustat;

y_0 – termenul luat drept bază de ajustare (este considerat acel termen al seriei empirice care se apropie în cea mai mare măsură de tendința medie trasată vizual);

t – variabila timp, respectiv numărul de intervale în sens negativ și pozitiv, față de baza de ajustare aleasă;

$$\bar{\Delta} - \text{sporul mediu} \left(\bar{\Delta} = \frac{\sum \Delta_{i-1}}{n} = \frac{\Delta_{n0}}{n} \right);$$

n – numărul de sporuri ale variației.

- *Metoda ritmului (indicielui) mediu.* Ajustarea pe baza ritmului mediu se aplică în cazul seriilor de timp ale căror termeni variază aproximativ după o *progresie geometrică*.

Termenii seriei ajustate se calculează după relația:

$$y_t = y_0 \cdot \bar{R}^t$$

în care: y_t – termenii seriei ajustate;

y_0 – termenul considerat ca bază de ajustare. Este ales acel termen al seriei empirice care se apropie cel mai mult, în reprezentarea grafică, de linia medie de evoluție (tendinței);

t – variabila timp determinată în raport cu baza de ajustare;

$\bar{R}$ – ritmul mediu ($\bar{R} = \sqrt[n]{\prod R_{i/i-1}} = \sqrt[n]{R_{n/0}}$);

n – numărul de sporuri ale variației.

- **Metoda analitică de ajustare** constă în *aproximarea tendinței de variație în timp cu ajutorul funcției de trend*, $y_t = f(t)$, în care „ t ” reprezintă variabila timp, iar „ y ” variabila în timp. Această metodă de ajustare ține seama de valorile empirice ale tuturor termenilor unei serii. Parametrii ecuației de trend se află, de regulă, cu ajutorul metodei celor mai mici pătrate.

- *Trendul liniar*

Ecuația de estimare a funcției de trend este definită de relația:

$$y_t = a + bt + e$$

Determinarea parametrilor funcției de trend are la bază **principiul celor mai mici pătrate**: $S = \sum e_i^2 = \sum (y_i - \hat{y}_i)^2 = \text{minim}$.

În funcția de trend variabila timp reprezintă, de regulă, o serie de numere consecutive, originea luându-se în centrul seriei și deci $\sum t_i = 0$. Ca urmare, relațiile de calcul pentru parametrii ecuației de trend sunt:

$$a = \frac{\sum y_i}{n} \text{ și } b = \frac{\sum t_i y_i}{\sum t_i^2}$$

- *Trendul parabolic*

Ecuția tendinței parabolice de gradul doi este:

$$y_t = a + b \cdot t + c \cdot t^2 + e$$

- *Trendul exponențial*

Ecuția trendului exponențial este:

$$y_t = a \cdot b^t$$

14.3. Ajustarea seriilor cronologice sezoniere

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp. 444-447)

Ajustarea seriilor cronologice cu variație sezonieră se bazează pe descompunerea seriei, pe de o parte, în componenta trend (f_t) și componenta ciclică C (când este cazul) și, pe de altă parte, în componenta variație sezonieră (S_t), considerând influența aleatoare nulă ($\sum e_t = 0$), adică:

$$y_t = f_t + S_t \text{ (pentru un model aditiv), respectiv}$$

$$y_t = f_t \cdot S_t \text{ (pentru un model multiplicativ).}$$

- Ajustarea prin medii mobile
- Coeficienți sezonieri
- Indici de sezonalitate

14.4. Criterii de alegere a celui mai bun model de trend

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp.448-450)

• Metoda grafică

- Se reprezintă grafic seria empirică
- Se ajustează vizual seria
- Se alege curba cea mai potrivită

- **Metoda celor mai mici pătrate**

$$\sum (y_i - y_t)^2 = \text{minim sau } \sum |y_i - y_t| = \text{minim}$$

- **Metoda coeficientului de variație**

$$v_{y_{it}} = \frac{\bar{d}_{y_{it}}}{y_i} \cdot 100, \quad v_{y_{it}} = \frac{s_{y_{it}}}{y_i} \cdot 100,$$

unde:

$$\bar{d}_{y_{it}} = \frac{\sum |y_i - y_t|}{n} \quad \text{abaterea medie liniară;}$$

$$\bar{y}_i = \frac{\sum y_i}{n} \quad \text{media;}$$

$$s_{y_{it}} = \sqrt{\frac{\sum (y_i - y_t)^2}{n}} \quad \text{abaterea medie pătratică.}$$

Se apreciază ca fiind cel mai potrivit model de trend *modelul pentru care coeficientul de variație este cel mai mic.*

14.5. Prognoza statistică

(Vezi Elisabeta Jaba, *Statistica*, Ediția a treia, Editura Economica, București, 2002, pp.450-460)

14.5.1. Probleme de prognoză

- Estimarea orizontului prognozei necesar dublării nivelului din prezent al unui fenomen;
- Estimarea orizontului prognozei necesar egalării nivelurilor unui fenomen din două zone diferite;
- Estimarea nivelului unui fenomen pentru un orizont de prognoză dat.

14.5.2. Orizontul prognozei

- **Orizontul prognozei în ipoteza evoluției după un model aditiv**

Funcția de trend, în condițiile progresiei aritmetice, ia forma:

$$y_n = y_0 + n \cdot \bar{\Delta}_y,$$

unde $\bar{\Delta}_y = \frac{y_n - y_0}{n}$, reprezintă sporul mediu de creștere, determinat pe baza datelor din perioada anterioară.

- *Orizontul prognozei necesar dublării nivelului din prezent al unui fenomen*
- *Orizontul prognozei necesar egalării nivelului de dezvoltare a două zone comparate*

- **Orizontul prognozei în ipoteza evoluției după un model multiplicativ**

Funcția de trend pentru un model multiplicativ este: $y_n = y_0 \cdot \bar{R}^n$,

unde: $\bar{R} = \sqrt[n]{\frac{y_n}{y_0}}$.

- *Orizontul prognozei necesar dublării nivelului din prezent al unui fenomen*
- *Orizontul prognozei necesar egalării nivelului de dezvoltare a două zone comparate*

14.5.3. Extrapolarea nivelului unui fenomen pentru un orizont de prognoză dat

- **Identificarea tendinței**
- **Analiza modelului de trend ales**
- **Extrapolarea**

14.5.4. Determinarea zonei de prognoză

Zona de variație, respectiv, intervalul de încredere pentru valoarea y_n , la diferite momente din viitor, se poate calcula după relația:

$$y_{(t)} - st_{\alpha/2} < y_n < y_{(t)} + st_{\alpha/2}$$

în care: $y_{(t)}$ este o variabilă y în timp;

t – variabila independentă – timp;

$t_{\alpha/2}$ – o valoare determinată din tabelele Student, pentru un coeficient de risc α (de regulă $\alpha = 0,05$, respectiv 5%) și ν grade de libertate;

s – abaterea medie pătratică la nivelul eșantionului.

$$s = \sqrt{s^2} = \sqrt{\frac{\sum (y_i - y_t)^2}{n - 2}}$$



14.6. Alegeți răspunsul corect

1. Trendul este componenta seriilor de timp care arată
 - a. variația medie pe termen lung
 - b. fluctuațiile regulate pe termen lung
 - c. fluctuațiile regulate într-o perioadă mai mică decât un an
2. Componenta ciclică este componenta seriilor de timp care arată
 - a. variația medie pe termen lung
 - b. fluctuațiile regulate pe termen lung
 - c. fluctuațiile regulate într-o perioadă mai mică decât un an
3. Componenta sezonieră este componenta seriilor de timp care arată
 - a. variația medie pe termen lung
 - b. fluctuațiile regulate pe termen lung
 - c. fluctuațiile regulate într-o perioadă mai mică decât un an
4. Componentele unei serii de timp pot fi combinate conform următoarelor modele:
 - a. modelul aditiv $y_t = f_t + C_t + S_t + e_t$
 - b. modelul multiplicativ $y_t = f_t \cdot C_t \cdot S_t \cdot e_t$
 - c. modelul mixt, în care sunt utilizate simultan modele aditive și multiplicative
5. Ajustarea seriilor cronologice se poate realiza prin
 - a. metoda grafică
 - b. metoda sporului mediu
 - c. metoda ritmului mediu
6. Ajustarea seriilor cronologice prin metoda sporului mediu se poate realiza când termenii variază aproximativ după o progresie
 - a. aritmetică
 - b. geometrică
 - c. aleatoare
7. Termenii seriei ajustate prin metoda sporului mediu se calculează după relația:
 - a. $y_t = y_0 + t \cdot \bar{R}$
 - b. $y_t = y_0 \cdot \bar{R}^t$
 - c. $y_t = y_0 + t \cdot \bar{\Delta}$

8. Ajustarea seriilor cronologice prin metoda ritmului mediu se poate realiza când termenii variază aproximativ după o progresie

- a. aritmetică
- b. geometrică
- c. aleatoare

9. Termenii seriei ajustate prin metoda ritmului mediu se calculează după relația:

- a. $y_t = y_0 + t \cdot \bar{R}$
- b. $y_t = y_0 \cdot \bar{R}'$
- c. $y_t = y_0 + t \cdot \bar{\Delta}$

10. Principiul celor mai mici pătrate presupune

- a. $S = \sum_i e_i^2 = \text{minim}$
- b. $S = \sum_i (y_i - y_{t_i})^2 = \text{minim}$
- c. $S = \sum_i (y_i - y_{t_i})^2 = \text{maxim}$

11. Relațiile de calcul pentru parametrii ecuației de trend când $\sum_i t_i = 0$ sunt

- a. $a = \frac{\sum_i y_i}{n}; b = \frac{\sum_i t_i y_i}{\sum_i t_i^2}$
- b. $a = \frac{\sum_i y_i \cdot \sum_i t_i^2 - \sum_i t_i \cdot \sum_i t_i \cdot y_i}{n \cdot \sum_i t_i^2 - \left(\sum_i t_i\right)^2}; b = \frac{n \cdot \sum_i t_i \cdot y_i - \sum_i t_i \cdot \sum_i y_i}{n \cdot \sum_i t_i^2 - \left(\sum_i t_i\right)^2}$
- c. $a = \frac{\sum_i y_i}{n} - b \cdot \frac{\sum_i t_i}{n}; b = \frac{n \cdot \sum_i t_i \cdot y_i - \sum_i t_i \cdot \sum_i y_i}{n \cdot \sum_i t_i^2 - \left(\sum_i t_i\right)^2}$

12. În cazul seriilor cronologice sezoniere, relațiile de calcul pentru parametrii ecuației de trend când $\sum_i t_i \neq 0$ sunt

$$a. \quad a = \frac{\sum_i y_i}{n}; b = \frac{\sum_i t_i y_i}{\sum_i t_i^2}$$

$$b. \quad a = \frac{\sum_i y_i \cdot \sum_i t_i^2 - \sum_i t_i \cdot \sum_i t_i \cdot y_i}{n \cdot \sum_i t_i^2 - \left(\sum_i t_i\right)^2}; b = \frac{n \cdot \sum_i t_i \cdot y_i - \sum_i t_i \cdot \sum_i y_i}{n \cdot \sum_i t_i^2 - \left(\sum_i t_i\right)^2}$$

$$c. \quad a = \frac{\sum_i y_i}{n} - b \cdot \frac{\sum_i t_i}{n}; b = \frac{n \cdot \sum_i t_i \cdot y_i - \sum_i t_i \cdot \sum_i y_i}{n \cdot \sum_i t_i^2 - \left(\sum_i t_i\right)^2}$$

13. Indicii de sezonalitate se calculează după relația

$$a. \quad i_t = \frac{y_t}{y_i}$$

$$b. \quad i_t = \frac{y_i}{y_t}$$

$$c. \quad i_t = \frac{y_i}{\bar{y}_i}$$

14. Coeficienții de sezonalitate se calculează

a. ca medie aritmetică a indicilor de sezonalitate

b. ca medie aritmetică a termenilor seriei ajustate

c. ca medie aritmetică a termenilor seriei observate

15. Criteriile de alegere a celui mai bun model de trend sunt

a. metoda grafică

b. metoda celor mai mici pătrate

c. metoda coeficientului de variație

16. Metoda grafică de alegere a celui mai bun model de trend presupune

a. reprezentarea grafică a seriei empirice și ajustarea vizuală a seriei după curbele care se apropie cel mai mult de linia medie de variație a fenomenului

b. respectarea principiului: $\sum_i (y_i - y_t)^2 = \text{minim}$

c. calculul coeficientului de variație, după relațiile:

$$v_{y_{III}} = \frac{\bar{d}_{y_{III}}}{\bar{y}_i} \cdot 100, v_{y_{III}} = \frac{s_{y_{III}}}{\bar{y}_i} \cdot 100$$

17. Metoda celor mai mici pătrate de alegere a celui mai bun model de trend presupune

a. reprezentarea grafică a seriei empirice și ajustarea vizuală a seriei după curbele care se apropie cel mai mult de linia medie de variație a fenomenului

b. respectarea principiului: $\sum_i (y_i - \bar{y}_i)^2 = \text{minim}$

c. calculul coeficientului de variație, după relațiile:

$$v_{y_{III}} = \frac{\bar{d}_{y_{III}}}{\bar{y}_i} \cdot 100, v_{y_{III}} = \frac{s_{y_{III}}}{\bar{y}_i} \cdot 100$$

18. Metoda coeficientului de variație de alegere a celui mai bun model de trend presupune

a. reprezentarea grafică a seriei empirice și ajustarea vizuală a seriei după curbele care se apropie cel mai mult de linia medie de variație a fenomenului

b. respectarea principiului: $\sum_i (y_i - \bar{y}_i)^2 = \text{minim}$

c. calculul coeficientului de variație, după relațiile:

$$v_{y_{III}} = \frac{\bar{d}_{y_{III}}}{\bar{y}_i} \cdot 100, v_{y_{III}} = \frac{s_{y_{III}}}{\bar{y}_i} \cdot 100$$

19. Prognoza statistică se poate realiza prin

a. metoda extrapolării tendinței de variație în timp

b. metode normative

c. metoda de regresie și corelație

20. Zona de prognoză, respectiv, intervalul de încredere, pentru valoarea y_{II} , se poate calcula după relația

a. $(y_i \pm s^2 \cdot t_{\alpha/2})$

b. $(\bar{y}_i \pm s \cdot t_{\alpha/2})$

c. $(y_i \pm s \cdot t_{\alpha/2})$



14.7. Recunoașteți exemplul corect

1. Evoluția cifrei de afaceri a unei firme, în perioada 1990-2000, poate fi aproximată printr-un model de forma: $y_t = 15 + 1,5t$. În această situație, se poate considera că

- a. cifra de afaceri a firmei a înregistrat o creștere medie anuală de 1,5 unități
- b. cifra de afaceri a firmei a înregistrat o reducere medie anuală de 1,5 unități
- c. cifra de afaceri a firmei a înregistrat o reducere medie anuală de 1,5 ori

2. Evoluția cifrei de afaceri a unei firme, în perioada 1990-2000, poate fi aproximată printr-un model de forma: $y_t = 15 - 1,5t$. În această situație, se poate considera că

- a. cifra de afaceri a firmei a înregistrat o creștere medie anuală de 1,5 unități
- b. cifra de afaceri a firmei a înregistrat o reducere medie anuală de 1,5 unități
- c. cifra de afaceri a firmei a înregistrat o reducere medie anuală de 1,5 ori

3. Evoluția unei variabile este exprimată prin următoarea ecuație de trend: $y_t = 2 + 5t$. Valoarea estimată pentru un orizont de prognoză $k=2$ este:

- a. 20
- b. 12
- c. 11

4. Pentru două modele de trend estimate pentru aceeași serie de timp s-au obținut următoarele rezultate privind valoarea coeficientului de variație: $v=3\%$ pentru modelul de trend de forma $y_t = a + bt$ și $v=15\%$ pentru modelul de forma $y_t = a + bt + ct^2$. Cel mai bun model de trend este:

- a. modelul liniar
- b. modelul parabolic
- c. modelul exponențial

5. Pentru două modele de trend estimate pentru aceeași serie de timp s-au obținut următoarele rezultate privind suma pătratelor abaterilor valorilor empirice de la valorile teoretice: $\sum (y_i - \hat{y}_i)^2 = 2,3$ pentru modelul de trend de

forma $y_i = a + bt$ și $\sum (y_i - y_{ti})^2 = 23$ pentru modelul de forma $y_i = a + bt + ct^2$. Cel mai bun model de trend este:

- a. modelul liniar
- b. modelul parabolic
- c. modelul exponențial

6. Pentru două modele de trend estimate pentru aceeași serie de timp s-au obținut următoarele rezultate privind suma abaterilor valorilor empirice de la valorile teoretice în valoare absolută: $\sum_i |y_i - y_{ti}| = 2,3$ pentru modelul de trend de forma $y_i = a + bt$ și $\sum_i |y_i - y_{ti}| = 1,2$ pentru modelul de forma $y_i = a + bt + ct^2$.

Cel mai bun model de trend este:

- a. modelul liniar
- b. modelul parabolic
- c. modelul exponențial



14.8. Răspunsuri corecte

14.6. Alegeți răspunsul corect

1. a) Trendul este componenta seriilor de timp care arată variația medie pe termen lung.
2. b) Componenta ciclică este componenta seriilor de timp care arată fluctuațiile regulate pe termen lung.
3. c) Componenta sezonieră este componenta seriilor de timp care arată fluctuațiile regulate într-o perioadă mai mică decât un an.
4. a), b) și c) Componentele unei serii de timp pot fi combinate conform următoarelor modele: modelul aditiv $y_t = f_t + C_t + S_t + e_t$, modelul multiplicativ $y_t = f_t \cdot C_t \cdot S_t \cdot e_t$, modelul mixt, în care sunt utilizate simultan modele aditive și multiplicative.
5. a), b) și c) Ajustarea seriilor cronologice se poate realiza prin metoda grafică, metoda sporului mediu, metoda ritmului mediu.
6. a) Ajustarea seriilor cronologice prin metoda sporului mediu se poate realiza când termenii variază aproximativ după o progresie aritmetică.
7. c) Termenii seriei ajustate prin metoda sporului mediu se calculează după relația: $y_t = y_0 + t \cdot \bar{\Delta}$.
8. b) Ajustarea seriilor cronologice prin metoda ritmului mediu se poate realiza când termenii variază aproximativ după o progresie geometrică.
9. b) Termenii seriei ajustate prin metoda ritmului mediu se calculează după relația: $y_t = y_0 \cdot \bar{R}^t$.
10. a) și b) Principiul celor mai mici pătrate presupune: $S = \sum_i e_i^2 = \text{minim}$,
 $S = \sum_i (y_i - y_{t_i})^2 = \text{minim}$.
11. a) Relațiile de calcul pentru parametrii ecuației de trend când $\sum_i t_i = 0$ sunt:

$$a = \frac{\sum_i y_i}{n}; b = \frac{\sum_i t_i y_i}{\sum_i t_i^2}.$$

12. b) și c) În cazul seriilor cronologice sezoniere, relațiile de calcul pentru parametrii ecuației de trend când $\sum_i t_i \neq 0$ sunt:

$$a = \frac{\sum_i y_i \cdot \sum_i t_i^2 - \sum_i t_i \cdot \sum_i t_i \cdot y_i}{n \cdot \sum_i t_i^2 - \left(\sum_i t_i\right)^2} \text{ sau } a = \frac{\sum_i y_i}{n} - b \cdot \frac{\sum_i t_i}{n};$$

$$b = \frac{n \cdot \sum_i t_i \cdot y_i - \sum_i t_i \cdot \sum_i y_i}{n \cdot \sum_i t_i^2 - \left(\sum_i t_i\right)^2}.$$

13. b) și c) Indicii de sezonalitate se calculează după relațiile: $i_t = \frac{y_t}{\bar{y}_t}$, respective $i_t = \frac{y_t}{\bar{y}_t}$.

14. a) Coeficienții de sezonalitate se calculează ca medie aritmetică a indicilor de sezonalitate.

15. a), b) și c) Criteriile de alegere a celui mai bun model de trend sunt metoda grafică, metoda celor mai mici pătrate și metoda coeficientului de variație.

16. a) Metoda grafică de alegere a celui mai bun model de trend presupune reprezentarea grafică a seriei empirice și ajustarea vizuală a seriei după curbele care se apropie cel mai mult de linia medie de variație a fenomenului.

17. b) Metoda celor mai mici pătrate de alegere a celui mai bun model de trend presupune respectarea principiului: $\sum_i (y_i - \hat{y}_i)^2 = \text{minim}$.

18. c) Metoda coeficientului de variație de alegere a celui mai bun model de trend presupune calculul coeficientului de variație, după relațiile:

$$v_{y_{it}} = \frac{\bar{d}_{y_{it}}}{\bar{y}_i} \cdot 100, v_{y_{it}} = \frac{s_{y_{it}}}{\bar{y}_i} \cdot 100.$$

19. a) și b) Prognoza statistică se poate realiza prin metoda extrapolării tendinței de variație în timp și metode normative.

20. c) Zona de prognoză, respectiv intervalul de încredere, pentru valoarea y_n , se poate calcula după relația $(y_i \pm s \cdot t_{\alpha/2})$.

14.7. Recunoașteți exemplul corect

1. a) Evoluția cifrei de afaceri a unei firme, în perioada 1990-2000, poate fi aproximată printr-un model de forma: $y_t = 15 + 1,5t$. În această situație, se poate considera că cifra de afaceri a firmei a înregistrat o creștere medie anuală de 1,5 unități.
2. b) Evoluția cifrei de afaceri a unei firme, în perioada 1990-2000, poate fi aproximată printr-un model de forma: $y_t = 15 - 1,5t$. În această situație, se poate considera că cifra de afaceri a firmei a înregistrat o reducere medie anuală de 1,5 unități.
3. b) Evoluția unei variabile este exprimată prin următoarea ecuație de trend: $y_t = 2 + 5t$. Valoarea estimată pentru un orizont de prognoză $k=2$ este $y_t = 12$.
4. a) Pentru două modele de trend estimate pentru aceeași serie de timp s-au obținut următoarele rezultate privind valoarea coeficientului de variație: $v=3\%$ pentru modelul de trend de forma $y_t = a + bt$ și $v=15\%$ pentru modelul de forma $y_t = a + bt + ct^2$. Cel mai bun model de trend este modelul liniar pentru care valoarea coeficientului de variație este minimă.
5. a) Pentru două modele de trend estimate pentru aceeași serie de timp, s-au obținut următoarele rezultate privind suma pătratelor abaterilor valorilor empirice de la valorile teoretice: $\sum (y_i - y_{ti})^2 = 2,3$ pentru modelul de trend de forma $y_t = a + bt$ și $\sum (y_i - y_{ti})^2 = 23$ pentru modelul de forma $y_t = a + bt + ct^2$. Cel mai bun model de trend este modelul liniar pentru care suma pătratelor abaterilor valorilor empirice de la valorile teoretice este minimă.
6. b) Pentru două modele de trend estimate pentru aceeași serie de timp s-au obținut următoarele rezultate privind suma abaterilor valorilor empirice de la valorile teoretice în valoare absolută: $\sum_i |y_i - y_{ti}| = 2,3$ pentru modelul de trend de forma $y_t = a + bt$ și $\sum_i |y_i - y_{ti}| = 1,2$ pentru modelul de forma $y_t = a + bt + ct^2$. Cel mai bun model de trend este modelul parabolic pentru care $\sum_i |y_i - y_{ti}| = \text{minim}$.

$$\text{acum: } y_x = \beta_0 + \beta_1 x, \quad y_x = b_0 + b_1 x$$



14.9. Probleme

14.9.1. Ajustarea seriilor de timp. Modelul de trend liniar

1. Asupra unui fenomen Y s-au înregistrat, pentru o perioadă de 7 ani, următoarele valori:

Anul	1	2	3	4	5	6	7
y_i	5	7	4	6	9	8	10

Se cere să se estimeze punctual parametrii ecuației de trend liniar, considerând $\sum_i t_i = 0$.

Când $\sum_i t_i = 0$, parametrii ecuației de trend liniar se estimează după relațiile:

$$a = \frac{\sum_i y_i}{n},$$

$$b = \frac{\sum_i t_i y_i}{\sum_i t_i^2}.$$

Estimarea parametrilor modelului de trend:

Elementele de calcul sunt prezentate în tabelul 14.1.

Estimațiile parametrilor modelului de trend sunt:

$$a = \frac{49}{7} = 7,$$

$$b = \frac{22}{28} = 0,786.$$

Ecuația de trend a variației fenomenului Y este de forma: $y_t = 7 + 0,786t$.

Interpretare:

Fenomenul Y a cunoscut, în cei 7 ani observați, o variație cu un trend

ascendent ($b > 0$). Nivelul fenomenului Y a crescut, în medie, cu 0,786 unități pe an.

Tabelul 14.1. Elemente de calcul

Anul	y_i	t_i	$t_i y_i$	t_i^2
1	5	-3	-15	9
2	7	-2	-14	4
3	4	-1	-4	1
4	6	0	0	0
5	9	+1	9	1
6	8	+2	16	4
7	10	+3	30	9
TOTAL	49	0	22	28

2. Asupra unui fenomen Y s-au înregistrat, pentru o perioadă de 7 ani, următoarele valori:

Anul	1	2	3	4	5	6	7
y_i	10	9	8	6	7	5	4

Se cere să se estimeze punctual parametrii ecuației de trend liniar, considerând $\sum_i t_i = 0$.

Când $\sum_i t_i = 0$, parametrii ecuației de trend liniar se estimează după relațiile:

$$a = \frac{\sum_i y_i}{n},$$

$$b = \frac{\sum_i t_i \cdot y_i}{\sum_i t_i^2}.$$

Estimarea parametrilor modelului de trend:

Elementele de calcul sunt prezentate în tabelul 14.2.

Estimațiile parametrilor modelului de trend sunt:

$$a = \frac{49}{7} = 7,$$

$$b = \frac{-27}{28} = -0,964.$$

Ecuția de trend a variației fenomenului Y este de forma: $y_t = 7 - 0,964t$.

Interpretare:

Fenomenul Y a cunoscut, în cei 7 ani observați, o evoluție cu un trend descendent ($b < 0$). Nivelul fenomenului Y s-a redus, în medie, cu 0,964 unități pe an.

Tabelul 14.2: Elemente de calcul

Anul	y_i	t_i	$t_i y_i$	t_i^2
1	10	-3	-30	9
2	9	-2	-18	4
3	8	-1	-8	1
4	6	0	0	0
5	7	+1	7	1
6	5	+2	10	4
7	4	+3	12	9
TOTAL	49	0	- 27	28

14.9.2. Prognoza statistică

1. În ipoteza că un fenomen urmează un trend liniar, $y_t = 10 + 3t$, se cere să se calculeze zona de prognoză pentru un orizont de prognoză $k=2$. Se cunosc următoarele date: $\sum_i (y_i - \hat{y}_i)^2 = 144$; $n=17$ ani; $\sum_i t_i = 0$. Se consideră un risc $\alpha=0,05$.

Zona de prognoză este definită de relația:

$$(y_t \pm s \cdot t_{\alpha/2; n-2}),$$

unde:

$$y_t = a + bt,$$

Pentru perioada de timp $n=17$ ani, când

$$\sum_i t_i = 0, \text{ nivelul fenomenului } Y \text{ pentru}$$

un orizont de prognoză $k=2$ este:

$$y_t = 10 + 3 \cdot 10 = 40$$

Când $\sum_i t_i = 0$, variabila timp ia valorile: $(0; \pm 1; \pm 2; \dots)$.

$$s = \sqrt{\frac{\sum_i (y_i - y_{t_i})^2}{n-2}},$$

$t_{\alpha/2; n-2}$ este valoarea statisticii t care urmează o distribuție de probabilitate Student cu $(n-2)$ grade de libertate, $t \sim t(n-2)$.

Estimația abaterii medii pătratice a erorilor este:

$$s = \sqrt{\frac{\sum_i (y_i - y_{t_i})^2}{n-2}} = \sqrt{\frac{144}{17-2}} = 3,1$$

Valoarea teoretică a statisticii test:

Valoarea teoretică a statisticii test t Student se citește în Tabela Student pentru un risc $\alpha=0,05$ și $v=17-2=15$ grade de libertate, $t_{0,025; 15}=2,131$.

α v	...	$t_{0.025}$	...
...	...	...	...
15	...	2,131	...
...	...	...	...

Zona de prognoză este:

$$(40 \pm 3,1 \cdot 2,131) = (33,39; 46,61)$$

Interpretare:

Dacă se mențin condițiile evoluției din trecut, se estimează că nivelul fenomenului Y va fi peste 2 ani acoperit de intervalul $(33,39; 46, 61)$. Se garantează acest rezultat cu o probabilitate de 0,95.

2. Ecuația de trend estimată la nivelul unui eșantion, $n=35$ ani, este definită de $y_t = 4 + 5t$. Se cere să se estimeze zona de prognoză pentru un orizont $k=2$.

Cunoaștem că $\sum_i e_i^2 = 49$, $\sum_i t_i = 0$ și considerăm un risc $\alpha = 0,05$.

Zona de prognoză este definită de relația:

$$(y_i \pm s \cdot z_{\alpha/2}),$$

unde:

$$y_i = a + b \cdot t,$$

Când $\sum_i t_i = 0$, variabila timp ia valorile: $(0; \pm 1; \pm 2; \dots)$.

$$s = \sqrt{\frac{\sum_i e_i^2}{n-2}} = \sqrt{\frac{\sum_i (y_i - y_{t_i})^2}{n-2}},$$

$z_{\alpha/2}$ este statistica normală standard Z pentru un risc α .

Pentru perioada de timp $n=35$ ani, când $\sum_i t_i = 0$, nivelul fenomenului Y pentru

un orizont de prognoză $k=2$ este:

$$y_i = 10 + 3 \cdot 20 = 70$$

Estimația abaterii medii pătratice a erorilor este:

$$s = \sqrt{\frac{49}{11-2}} = 1,22$$

Valoarea teoretică a statisticii test:

Valoarea teoretică a statisticii Z se citește în Tabela Gauss-Laplace pentru un risc $\alpha=0,05$, $z=1,96$.

z	...	0.06	...
:			
1.9	...	0,475	...
:			

Zona de prognoză:

$$(70 \pm 1,22 \cdot 1,96) = (67,61; 72,39)$$

Interpretare: Dacă se mențin condițiile evoluției din trecut, se estimează că nivelul fenomenului Y va fi peste 2 ani acoperit de intervalul: $(67,61; 72,39)$. Se garantează acest rezultat cu o probabilitate de 0,95.

14.9.3. Ajustarea seriilor de timp cu caracter sezonier

1. Cererea unui produs a înregistrat, pe parcursul a doi ani, o evoluție definită de ecuația de trend $y_t = 3 + 16t$ și următorii indici de sezonalitate:

Trimestrul	Anul 1	Anul 2
I	1,6	1,4
II	1,2	1,0
III	0,9	0,8
IV	0,9	0,7

Se cere să se estimeze nivelul fenomenului Y pentru trimestrul patru al anului trei.

Extrapolarea trendului presupune calculul valorii Y pentru $t=n+k$, după relația:

$$y_{t=n+k} = a + b \cdot (t = n + k)$$

Corectarea valorii prognozate cu valoarea coeficientului de sezonalitate corespunzător trimestrului patru (S_j):

$$y_t^{(4)} = y_{t=n+k} \cdot S_j, \text{ unde}$$

S_j este media aritmetică a indicilor de sezonalitate pentru trimestrul patru corespunzător celor doi ani observați:

$$S_j = \frac{i_4 + i_8}{2}$$

Extrapolarea trendului:

– în trimestrul patru al anului 3 valoarea lui t este $t=8+4=12$. Nivelul fenomenului Y pentru acest orizont este:

$$y_t = 3 + 16 \cdot 12 = 195$$

Coeficientul de sezonalitate pentru trimestrul patru corespunzător celor doi ani considerați este:

$$S_j = \frac{0,9 + 0,7}{2} = 0,8$$

Corectarea valorii prognozate cu valoarea coeficientului de sezonalitate corespunzător trimestrului patru (S_j):

$$y_t^{(4)} = 195 \cdot 0,8 = 156$$

Interpretare:

Dacă se mențin condițiile evoluției din trecut, se estimează că nivelul fenomenului Y va fi, în trimestrul patru al anului următor celor observați, de 156.

2. Cererea unui produs a înregistrat, pe parcursul a trei ani, o evoluție definită de ecuația de trend $y_t = 4 + 7,5t$ și următorii indici de sezonality:

Trimestrul	Anul 1	Anul 2	Anul 3
I	1,6	1,4	1,7
II	1,2	1,0	1,3
III	0,9	0,8	0,7
IV	0,9	1,1	1,0

Se cere să se estimeze nivelul fenomenului Y pentru trimestrul trei al anului patru.

Extrapolarea trendului presupune calculul valorii Y pentru $t=n+k$, după relația:

$$y_{t=n+k} = a + b \cdot (t = n + k)$$

Corectarea valorii prognozate cu valoarea coeficientului de sezonality corespunzător trimestrului trei (S_j):

$$y_t^{(3)} = y_{t=n+k} \cdot S_j, \text{ unde}$$

S_j este media aritmetică a indicilor de sezonality pentru trimestrul trei corespunzător celor trei ani considerați:

$$S_j = \frac{i_3 + i_7 + i_{11}}{3}$$

Extrapolarea trendului:

– în trimestrul trei al anului 4 valoarea lui t este $t=12+3=15$. Nivelul fenomenului Y pentru acest orizont este:

$$y_t = 4 + 7,5 \cdot 15 = 116,5$$

Coeficientul de sezonality pentru trimestrul trei corespunzător celor trei ani considerați este:

$$S_j = \frac{0,9 + 0,8 + 0,7}{3} = 0,8$$

Corectarea valorii prognozate cu valoarea coeficientului de sezonality corespunzător trimestrului trei (S_j):

$$y_t^{(3)} = 116,5 \cdot 0,8 = 93,2$$

Interpretare:

Dacă se mențin condițiile evoluției din trecut, se estimează că nivelul fenomenului Y va fi, în trimestrul trei al anului următor celor doi ani observați, de 93,2.

3. O firmă înregistrează valoarea vânzărilor trimestriale (miliarde lei) pentru o perioadă de 2 ani. În urma prelucrării datelor, s-au obținut următoarele rezultate:

Anul	Trimestrul	y_i	$y_{ti}=5,36+0,14ti$
2003	I	5	5,5
2003	II	8	5,64
2003	III	4	5,78
2003	IV	3	5,92
2004	I	7	6,06
2004	II	10	6,2
2004	III	6	6,34
2004	IV	5	6,48
TOTAL	-	44	-

Se cere să se estimeze valoarea vânzărilor firmei pentru trimestrul trei al anului 2005.

Extrapolarea trendului presupune calculul valorii Y pentru $t=n+k$, după relația:

$$y_{t=n+k} = a + b \cdot (t = n + k)$$

Corectarea valorii prognozate cu valoarea coeficientului de sezonalitate corespunzător trimestrului trei (S_j):

$$y_t^{(3)} = y_{t=n+k} \cdot S_j, \text{ unde}$$

S_j este media aritmetică a indicilor de sezonalitate pentru trimestrul trei corespunzător celor doi ani considerați:

$$S_j = \frac{i_3 + i_7}{2}.$$

Indicii de sezonalitate se calculează ca raport între valorile reale (y_i) și valorile teoretice (y_{ti}), după relația:

Extrapolarea trendului:

– în trimestrul trei al anului 4 valoarea lui t este $t=8+3=11$. Nivelul fenomenului Y pentru acest orizont este:

$$y_t = 5,36 + 0,14 \cdot 11 = 6,9$$

Indicii de sezonalitate pentru trimestrul trei ai celor doi ani observați sunt:

– anul 2003:

$$i_3 = \frac{4}{5,78} = 0,692$$

– anul 2004:

$$i_7 = \frac{6}{6,34} = 0,946$$

Coeficientul de sezonalitate pentru trimestrul trei corespunzător celor doi ani considerați este:

$$S_j = \frac{0,692 + 0,946}{2} = 0,819$$

$$i_i = \frac{y_i}{y_t}$$

Corectarea valorii prognozate cu valoarea coeficientului de sezonalitate corespunzător trimestrului trei (S_j):

$$y_t^{(3)} = 6,9 \cdot 0,819 = 5,65$$

Interpretare:

Dacă se mențin condițiile evoluției din trecut, se estimează că nivelul fenomenului Y va fi, în trimestrul trei al anului 2005, de 5,65 mld.lei.

4. O firmă înregistrează valoarea vânzărilor trimestriale (miliarde lei) pentru o perioadă de 3 ani. În urma prelucrării datelor, s-au obținut următoarele rezultate:

Anul	Trimestrul	y_i	$yt_i=4,62+0,39ti$
2002	I	5	5,01
2002	II	8	5,4
2002	III	4	5,79
2002	IV	3	6,18
2003	I	7	6,57
2003	II	10	6,96
2003	III	6	7,35
2003	IV	5	7,74
2004	I	9	8,13
2004	II	14	8,52
2004	III	8	8,91
2004	IV	7	9,3
TOTAL	-	86	-

Se cere să se estimeze valoarea vânzărilor firmei pentru trimestrul doi al anului 2005.

Extrapolarea trendului presupune calculul valorii Y pentru $t=n+k$, după relația:

$$y_{t=n+k} = a + b \cdot (t = n + k)$$

Extrapolarea trendului:

- în trimestrul trei al anului 4 valoarea lui t este $t=12+2=14$. Nivelul fenomenului Y pentru acest orizont este:

$$y_t = 4,62 + 0,39 \cdot 14 = 10,08$$

Corectarea valorii prognozate cu valoarea coeficientului de sezonabilitate corespunzător trimestrului doi (S_j):

$$y_t^{(2)} = y_{t=n+k} \cdot S_j, \text{ unde}$$

S_j este media aritmetică a indicilor de sezonabilitate pentru trimestrului doi corespunzător celor trei ani considerați:

$$S_j = \frac{i_2 + i_6 + i_{10}}{3}.$$

Indicii de sezonabilitate se calculează ca raport între valorile reale (y_i) și valorile teoretice (y_{ti}), după relația:

$$i_{t_i} = \frac{y_i}{y_{t_i}}.$$

Indicii de sezonabilitate pentru trimestrului doi ai celor trei ani observați sunt:

– anul 2002:

$$i_2 = \frac{8}{5,4} = 1,481$$

– anul 2003:

$$i_6 = \frac{10}{6,96} = 1,437$$

– anul 2004:

$$i_{10} = \frac{14}{8,52} = 1,643$$

Coeficientul de sezonabilitate pentru trimestrului doi corespunzător celor trei ani considerați este:

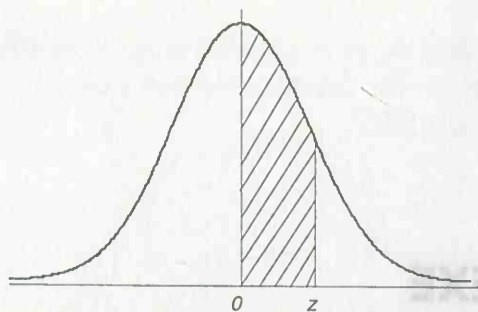
$$S_j = \frac{1,481 + 1,437 + 1,643}{3} = 1,52$$

Corectarea valorii prognozate cu valoarea coeficientului de sezonabilitate corespunzător trimestrului doi (S_j):

$$y_t^{(3)} = 10,08 \cdot 1,52 = 15,32$$

Interpretare:

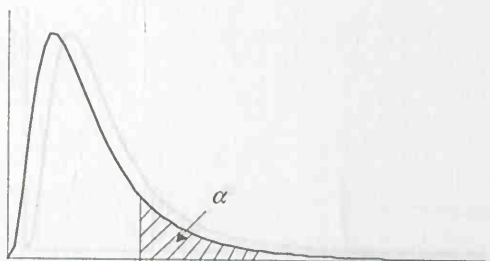
Dacă se mențin condițiile evoluției din trecut, se estimează că nivelul fenomenului Y va fi, în trimestrului doi al anului 2005, de 15,32 mld.lei.



Valorile funcției lui Laplace:

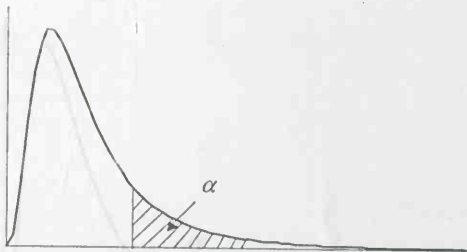
$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}} dt$$

[illegible]

Valorile distribuției F, $\alpha = 0,1$

	1	2	3	4	5	6	7	8	9
1	39.864	49.500	53.593	55.833	57.240	58.204	58.906	59.439	59.857
2	8.526	9.000	9.162	9.243	9.293	9.326	9.349	9.367	9.381
3	5.538	5.462	5.391	5.343	5.309	5.285	5.266	5.252	5.240
4	4.545	4.325	4.191	4.107	4.051	4.010	3.979	3.955	3.936
5	4.060	3.780	3.619	3.520	3.453	3.405	3.368	3.339	3.316
6	3.776	3.463	3.289	3.181	3.108	3.055	3.014	2.983	2.958
7	3.589	3.257	3.074	2.961	2.883	2.827	2.785	2.752	2.725
8	3.458	3.113	2.924	2.806	2.726	2.668	2.624	2.589	2.561
9	3.360	3.006	2.813	2.693	2.611	2.551	2.505	2.469	2.440
10	3.285	2.924	2.728	2.605	2.522	2.461	2.414	2.377	2.347
11	3.225	2.860	2.660	2.536	2.451	2.389	2.342	2.304	2.274
12	3.177	2.807	2.606	2.480	2.394	2.331	2.283	2.245	2.214
13	3.136	2.763	2.560	2.434	2.347	2.283	2.234	2.195	2.164
14	3.102	2.726	2.522	2.395	2.307	2.243	2.193	2.154	2.122
15	3.073	2.695	2.490	2.361	2.273	2.208	2.158	2.119	2.086
16	3.048	2.668	2.462	2.333	2.244	2.178	2.128	2.088	2.055
17	3.026	2.645	2.437	2.308	2.218	2.152	2.102	2.061	2.028
18	3.007	2.624	2.416	2.286	2.196	2.130	2.079	2.038	2.005
19	2.990	2.606	2.397	2.266	2.176	2.109	2.058	2.017	1.984
20	2.975	2.589	2.380	2.249	2.158	2.091	2.040	1.999	1.965
21	2.961	2.575	2.365	2.233	2.142	2.075	2.023	1.982	1.948
22	2.949	2.561	2.351	2.219	2.128	2.060	2.008	1.967	1.933
23	2.937	2.549	2.339	2.207	2.115	2.047	1.995	1.953	1.919
24	2.927	2.538	2.327	2.195	2.103	2.035	1.983	1.941	1.906
25	2.918	2.528	2.317	2.184	2.092	2.024	1.971	1.929	1.895
26	2.909	2.519	2.307	2.174	2.082	2.014	1.961	1.919	1.884
27	2.901	2.511	2.299	2.165	2.073	2.005	1.952	1.909	1.874
28	2.894	2.503	2.291	2.157	2.064	1.996	1.943	1.900	1.865
29	2.887	2.495	2.283	2.149	2.057	1.988	1.935	1.892	1.857
30	2.881	2.489	2.276	2.142	2.049	1.980	1.927	1.884	1.849
40	2.835	2.440	2.226	2.091	1.997	1.927	1.873	1.829	1.793
60	2.791	2.393	2.177	2.041	1.946	1.875	1.819	1.775	1.738
120	2.748	2.347	2.130	1.992	1.896	1.824	1.767	1.722	1.684
∞	2.706	2.303	2.084	1.945	1.847	1.774	1.717	1.670	1.632

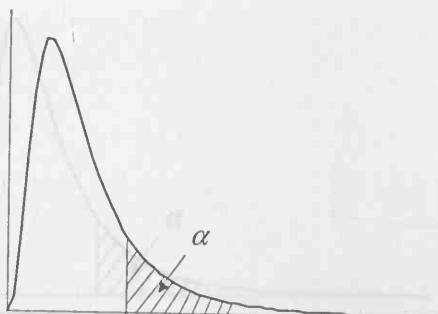
	10	12	15	20	24	30	40	60	120	∞
1	60.195	60.705	61.220	61.740	62.002	62.265	62.529	62.794	63.061	63.328
2	9.392	9.408	9.425	9.441	9.450	9.458	9.466	9.475	9.483	9.491
3	5.230	5.216	5.200	5.184	5.176	5.168	5.160	5.151	5.143	5.134
4	3.920	3.896	3.870	3.844	3.831	3.817	3.804	3.790	3.775	3.761
5	3.297	3.268	3.238	3.207	3.191	3.174	3.157	3.140	3.123	3.105
6	2.937	2.905	2.871	2.836	2.818	2.800	2.781	2.762	2.742	2.722
7	2.703	2.668	2.632	2.595	2.575	2.555	2.535	2.514	2.493	2.471
8	2.538	2.502	2.464	2.425	2.404	2.383	2.361	2.339	2.316	2.293
9	2.416	2.379	2.340	2.298	2.277	2.255	2.232	2.208	2.184	2.159
10	2.323	2.284	2.244	2.201	2.178	2.155	2.132	2.107	2.082	2.055
11	2.248	2.209	2.167	2.123	2.100	2.076	2.052	2.026	2.000	1.972
12	2.188	2.147	2.105	2.060	2.036	2.011	1.986	1.960	1.932	1.904
13	2.138	2.097	2.053	2.007	1.983	1.958	1.931	1.904	1.876	1.846
14	2.095	2.054	2.010	1.962	1.938	1.912	1.885	1.857	1.828	1.797
15	2.059	2.017	1.972	1.924	1.899	1.873	1.845	1.817	1.787	1.755
16	2.028	1.985	1.940	1.891	1.866	1.839	1.811	1.782	1.751	1.718
17	2.001	1.958	1.912	1.862	1.836	1.809	1.781	1.751	1.719	1.686
18	1.977	1.933	1.887	1.837	1.810	1.783	1.754	1.723	1.691	1.657
19	1.956	1.912	1.865	1.814	1.787	1.759	1.730	1.699	1.666	1.631
20	1.937	1.892	1.845	1.794	1.767	1.738	1.708	1.677	1.643	1.607
21	1.920	1.875	1.827	1.776	1.748	1.719	1.689	1.657	1.623	1.586
22	1.904	1.859	1.811	1.759	1.731	1.702	1.671	1.639	1.604	1.567
23	1.890	1.845	1.796	1.744	1.716	1.686	1.655	1.622	1.587	1.549
24	1.877	1.832	1.783	1.730	1.702	1.672	1.641	1.607	1.571	1.533
25	1.866	1.820	1.771	1.718	1.689	1.659	1.627	1.593	1.557	1.518
26	1.855	1.809	1.760	1.706	1.677	1.647	1.615	1.581	1.544	1.504
27	1.845	1.799	1.749	1.695	1.666	1.636	1.603	1.569	1.531	1.491
28	1.836	1.790	1.740	1.685	1.656	1.625	1.592	1.558	1.520	1.478
29	1.827	1.781	1.731	1.676	1.647	1.616	1.583	1.547	1.509	1.467
30	1.819	1.773	1.722	1.667	1.638	1.606	1.573	1.538	1.499	1.456
40	1.763	1.715	1.662	1.605	1.574	1.541	1.506	1.467	1.425	1.377
60	1.707	1.657	1.603	1.543	1.511	1.476	1.437	1.395	1.348	1.291
120	1.652	1.601	1.545	1.482	1.447	1.409	1.368	1.320	1.265	1.193
∞	1.599	1.546	1.487	1.421	1.383	1.342	1.295	1.240	1.169	1.000



Valorile distribuției F, $\alpha = 0,05$

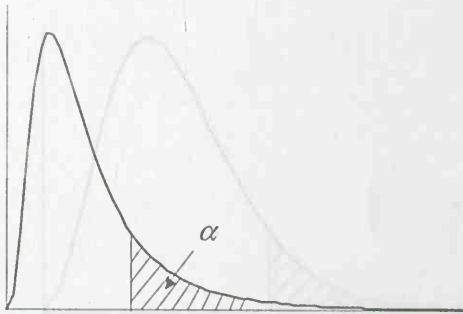
	1	2	3	4	5	6	7	8	9
1	161.446	199.499	215.707	224.583	230.160	233.988	236.767	238.884	240.543
2	18.513	19.000	19.164	19.247	19.296	19.329	19.353	19.371	19.385
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999
5	6.606	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677
8	5.318	4.459	4.066	3.838	3.688	3.581	3.500	3.438	3.388
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420	2.366
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397	2.342
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375	2.320
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321	2.265
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305	2.250
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291	2.236
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278	2.223
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040
120	3.920	3.072	2.680	2.447	2.290	2.175	2.087	2.016	1.959
∞	3.841	2.996	2.605	2.372	2.214	2.099	2.010	1.938	1.880

10	12	15	20	24	30	40	60	120	∞
241.882	243.905	245.949	248.016	249.052	250.096	251.144	252.196	253.254	254.317
19.396	19.412	19.429	19.446	19.454	19.463	19.471	19.479	19.487	19.496
8.785	8.745	8.703	8.660	8.638	8.617	8.594	8.572	8.549	8.526
5.964	5.912	5.858	5.803	5.774	5.746	5.717	5.688	5.658	5.628
4.735	4.678	4.619	4.558	4.527	4.496	4.464	4.431	4.398	4.365
4.060	4.000	3.938	3.874	3.841	3.808	3.774	3.740	3.705	3.669
3.637	3.575	3.511	3.445	3.410	3.376	3.340	3.304	3.267	3.230
3.347	3.284	3.218	3.150	3.115	3.079	3.043	3.005	2.967	2.928
3.137	3.073	3.006	2.936	2.900	2.864	2.826	2.787	2.748	2.707
2.978	2.913	2.845	2.774	2.737	2.700	2.661	2.621	2.580	2.538
2.854	2.788	2.719	2.646	2.609	2.570	2.531	2.490	2.448	2.404
2.753	2.687	2.617	2.544	2.505	2.466	2.426	2.384	2.341	2.296
2.671	2.604	2.533	2.459	2.420	2.380	2.339	2.297	2.252	2.206
2.602	2.534	2.463	2.388	2.349	2.308	2.266	2.223	2.178	2.131
2.544	2.475	2.403	2.328	2.288	2.247	2.204	2.160	2.114	2.066
2.494	2.425	2.352	2.276	2.235	2.194	2.151	2.106	2.059	2.010
2.450	2.381	2.308	2.230	2.190	2.148	2.104	2.058	2.011	1.960
2.412	2.342	2.269	2.191	2.150	2.107	2.063	2.017	1.968	1.917
2.378	2.308	2.234	2.155	2.114	2.071	2.026	1.980	1.930	1.878
2.348	2.278	2.203	2.124	2.082	2.039	1.994	1.946	1.896	1.843
2.321	2.250	2.176	2.096	2.054	2.010	1.965	1.916	1.866	1.812
2.297	2.226	2.151	2.071	2.028	1.984	1.938	1.889	1.838	1.783
2.275	2.204	2.128	2.048	2.005	1.961	1.914	1.865	1.813	1.757
2.255	2.183	2.108	2.027	1.984	1.939	1.892	1.842	1.790	1.733
2.236	2.165	2.089	2.007	1.964	1.919	1.872	1.822	1.768	1.711
2.220	2.148	2.072	1.990	1.946	1.901	1.853	1.803	1.749	1.691
2.204	2.132	2.056	1.974	1.930	1.884	1.836	1.785	1.731	1.672
2.190	2.118	2.041	1.959	1.915	1.869	1.820	1.769	1.714	1.654
2.177	2.104	2.027	1.945	1.901	1.854	1.806	1.754	1.698	1.638
2.165	2.092	2.015	1.932	1.887	1.841	1.792	1.740	1.683	1.622
2.077	2.003	1.924	1.839	1.793	1.744	1.693	1.637	1.577	1.509
1.993	1.917	1.836	1.748	1.700	1.649	1.594	1.534	1.467	1.389
1.910	1.834	1.750	1.659	1.608	1.554	1.495	1.429	1.352	1.254
1.831	1.752	1.666	1.571	1.517	1.459	1.394	1.318	1.221	1.000

Valorile distribuției F, $\alpha = 0,025$

	1	2	3	4	5	6	7	8	9
1	647.793	799.482	864.151	899.599	921.835	937.114	948.203	956.643	963.279
2	38.506	39.000	39.166	39.248	39.298	39.331	39.356	39.373	39.387
3	17.443	16.044	15.439	15.101	14.885	14.735	14.624	14.540	14.473
4	12.218	10.649	9.979	9.604	9.364	9.197	9.074	8.980	8.905
5	10.007	8.434	7.764	7.388	7.146	6.978	6.853	6.757	6.681
6	8.813	7.260	6.599	6.227	5.988	5.820	5.695	5.600	5.523
7	8.073	6.542	5.890	5.523	5.285	5.119	4.995	4.899	4.823
8	7.571	6.059	5.416	5.053	4.817	4.652	4.529	4.433	4.357
9	7.209	5.715	5.078	4.718	4.484	4.320	4.197	4.102	4.026
10	6.937	5.456	4.826	4.468	4.236	4.072	3.950	3.855	3.779
11	6.724	5.256	4.630	4.275	4.044	3.881	3.759	3.664	3.588
12	6.554	5.096	4.474	4.121	3.891	3.728	3.607	3.512	3.436
13	6.414	4.965	4.347	3.996	3.767	3.604	3.483	3.388	3.312
14	6.298	4.857	4.242	3.892	3.663	3.501	3.380	3.285	3.209
15	6.200	4.765	4.153	3.804	3.576	3.415	3.293	3.199	3.123
16	6.115	4.687	4.077	3.729	3.502	3.341	3.219	3.125	3.049
17	6.042	4.619	4.011	3.665	3.438	3.277	3.156	3.061	2.985
18	5.978	4.560	3.954	3.608	3.382	3.221	3.100	3.005	2.929
19	5.922	4.508	3.903	3.559	3.333	3.172	3.051	2.956	2.880
20	5.871	4.461	3.859	3.515	3.289	3.128	3.007	2.913	2.837
21	5.827	4.420	3.819	3.475	3.250	3.090	2.969	2.874	2.798
22	5.786	4.383	3.783	3.440	3.215	3.055	2.934	2.839	2.763
23	5.750	4.349	3.750	3.408	3.183	3.023	2.902	2.808	2.731
24	5.717	4.319	3.721	3.379	3.155	2.995	2.874	2.779	2.703
25	5.686	4.291	3.694	3.353	3.129	2.969	2.848	2.753	2.677
26	5.659	4.265	3.670	3.329	3.105	2.945	2.824	2.729	2.653
27	5.633	4.242	3.647	3.307	3.083	2.923	2.802	2.707	2.631
28	5.610	4.221	3.626	3.286	3.063	2.903	2.782	2.687	2.611
29	5.588	4.201	3.607	3.267	3.044	2.884	2.763	2.669	2.592
30	5.568	4.182	3.589	3.250	3.026	2.867	2.746	2.651	2.575
40	5.424	4.051	3.463	3.126	2.904	2.744	2.624	2.529	2.452
60	5.286	3.925	3.343	3.008	2.786	2.627	2.507	2.412	2.334
120	5.152	3.805	3.227	2.894	2.674	2.515	2.395	2.299	2.222
∞	5.024	3.689	3.116	2.786	2.567	2.408	2.288	2.192	2.114

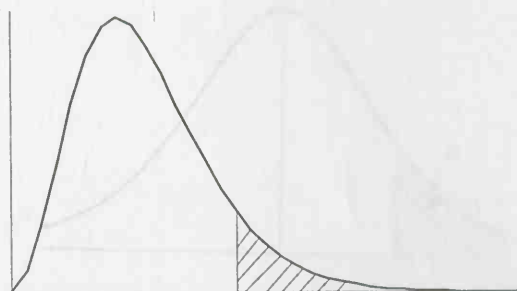
	10	12	15	20	24	30	40	60	120	∞
1	968.63	976.72	984.87	993.08	997.27	1001	1005	1009	1014	1018
2	39.398	39.415	39.431	39.448	39.457	39.465	39.473	39.481	39.489	39.498
3	14.419	14.337	14.253	14.167	14.124	14.081	14.036	13.992	13.947	13.902
4	8.844	8.751	8.657	8.560	8.511	8.461	8.411	8.360	8.309	8.257
5	6.619	6.525	6.428	6.329	6.278	6.227	6.175	6.123	6.069	6.015
6	5.461	5.366	5.269	5.168	5.117	5.065	5.012	4.959	4.904	4.849
7	4.761	4.666	4.568	4.467	4.415	4.362	4.309	4.254	4.199	4.142
8	4.295	4.200	4.101	3.999	3.947	3.894	3.840	3.784	3.728	3.670
9	3.964	3.868	3.769	3.667	3.614	3.560	3.505	3.449	3.392	3.333
10	3.717	3.621	3.522	3.419	3.365	3.311	3.255	3.198	3.140	3.080
11	3.526	3.430	3.330	3.226	3.173	3.118	3.061	3.004	2.944	2.883
12	3.374	3.277	3.177	3.073	3.019	2.963	2.906	2.848	2.787	2.725
13	3.250	3.153	3.053	2.948	2.893	2.837	2.780	2.720	2.659	2.595
14	3.147	3.050	2.949	2.844	2.789	2.732	2.674	2.614	2.552	2.487
15	3.060	2.963	2.862	2.756	2.701	2.644	2.585	2.524	2.461	2.395
16	2.986	2.889	2.788	2.681	2.625	2.568	2.509	2.447	2.383	2.316
17	2.922	2.825	2.723	2.616	2.560	2.502	2.442	2.380	2.315	2.247
18	2.866	2.769	2.667	2.559	2.503	2.445	2.384	2.321	2.256	2.187
19	2.817	2.720	2.617	2.509	2.452	2.394	2.333	2.270	2.203	2.133
20	2.774	2.676	2.573	2.464	2.408	2.349	2.287	2.223	2.156	2.085
21	2.735	2.637	2.534	2.425	2.368	2.308	2.246	2.182	2.114	2.042
22	2.700	2.602	2.498	2.389	2.332	2.272	2.210	2.145	2.076	2.003
23	2.668	2.570	2.466	2.357	2.299	2.239	2.176	2.111	2.041	1.968
24	2.640	2.541	2.437	2.327	2.269	2.209	2.146	2.080	2.010	1.935
25	2.613	2.515	2.411	2.300	2.242	2.182	2.118	2.052	1.981	1.906
26	2.590	2.491	2.387	2.276	2.217	2.157	2.093	2.026	1.954	1.878
27	2.568	2.469	2.364	2.253	2.195	2.133	2.069	2.002	1.930	1.853
28	2.547	2.448	2.344	2.232	2.174	2.112	2.048	1.980	1.907	1.829
29	2.529	2.430	2.325	2.213	2.154	2.092	2.028	1.959	1.886	1.807
30	2.511	2.412	2.307	2.195	2.136	2.074	2.009	1.940	1.866	1.787
40	2.388	2.288	2.182	2.068	2.007	1.943	1.875	1.803	1.724	1.637
60	2.270	2.169	2.061	1.944	1.882	1.815	1.744	1.667	1.581	1.482
120	2.157	2.055	1.945	1.825	1.760	1.690	1.614	1.530	1.433	1.310
∞	2.048	1.945	1.833	1.708	1.640	1.566	1.484	1.388	1.268	1.000



Valorile distribuției F, $\alpha = 0,01$

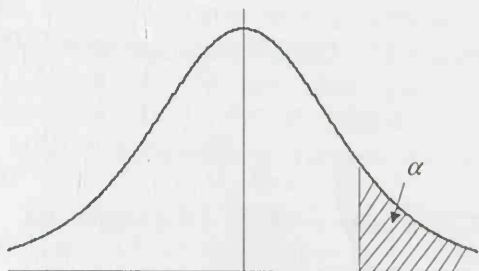
	1	2	3	4	5	6	7	8	9
1	4052	4999	5403	5624	5763	5858	5928	5980	6022
2	98.502	99.000	99.164	99.251	99.302	99.331	99.357	99.375	99.390
3	34.116	30.816	29.457	28.710	28.237	27.911	27.671	27.489	27.345
4	21.198	18.000	16.694	15.977	15.522	15.207	14.976	14.799	14.659
5	16.258	13.274	12.060	11.392	10.967	10.672	10.456	10.289	10.158
6	13.745	10.925	9.780	9.148	8.746	8.466	8.260	8.102	7.976
7	12.246	9.547	8.451	7.847	7.460	7.191	6.993	6.840	6.719
8	11.259	8.649	7.591	7.006	6.632	6.371	6.178	6.029	5.911
9	10.562	8.022	6.992	6.422	6.057	5.802	5.613	5.467	5.351
10	10.044	7.559	6.552	5.994	5.636	5.386	5.200	5.057	4.942
11	9.646	7.206	6.217	5.668	5.316	5.069	4.886	4.744	4.632
12	9.330	6.927	5.953	5.412	5.064	4.821	4.640	4.499	4.388
13	9.074	6.701	5.739	5.205	4.862	4.620	4.441	4.302	4.191
14	8.862	6.515	5.564	5.035	4.695	4.456	4.278	4.140	4.030
15	8.683	6.359	5.417	4.893	4.556	4.318	4.142	4.004	3.895
16	8.531	6.226	5.292	4.773	4.437	4.202	4.026	3.890	3.780
17	8.400	6.112	5.185	4.669	4.336	4.101	3.927	3.791	3.682
18	8.285	6.013	5.092	4.579	4.248	4.015	3.841	3.705	3.597
19	8.185	5.926	5.010	4.500	4.171	3.939	3.765	3.631	3.523
20	8.096	5.849	4.938	4.431	4.103	3.871	3.699	3.564	3.457
21	8.017	5.780	4.874	4.369	4.042	3.812	3.640	3.506	3.398
22	7.945	5.719	4.817	4.313	3.988	3.758	3.587	3.453	3.346
23	7.881	5.664	4.765	4.264	3.939	3.710	3.539	3.406	3.299
24	7.823	5.614	4.718	4.218	3.895	3.667	3.496	3.363	3.256
25	7.770	5.568	4.675	4.177	3.855	3.627	3.457	3.324	3.217
26	7.721	5.526	4.637	4.140	3.818	3.591	3.421	3.288	3.182
27	7.677	5.488	4.601	4.106	3.785	3.558	3.388	3.256	3.149
28	7.636	5.453	4.568	4.074	3.754	3.528	3.358	3.226	3.120
29	7.598	5.420	4.538	4.045	3.725	3.499	3.330	3.198	3.092
30	7.562	5.390	4.510	4.018	3.699	3.473	3.305	3.173	3.067
40	7.314	5.178	4.313	3.828	3.514	3.291	3.124	2.993	2.888
60	7.077	4.977	4.126	3.649	3.339	3.119	2.953	2.823	2.718
120	6.851	4.787	3.949	3.480	3.174	2.956	2.792	2.663	2.559
∞	6.635	4.605	3.782	3.319	3.017	2.802	2.639	2.511	2.407

	10	12	15	20	24	30	40	60	120	∞
1	6055	6106	6156	6208	6234	6260	6286	6312	6339	6365
2	99.397	99.419	99.433	99.448	99.455	99.466	99.477	99.484	99.491	99.499
3	27.228	27.052	26.872	26.690	26.597	26.504	26.411	26.316	26.221	26.125
4	14.546	14.374	14.198	14.019	13.929	13.838	13.745	13.652	13.558	13.463
5	10.051	9.888	9.722	9.553	9.466	9.379	9.291	9.202	9.112	9.020
6	7.874	7.718	7.559	7.396	7.313	7.229	7.143	7.057	6.969	6.880
7	6.620	6.469	6.314	6.155	6.074	5.992	5.908	5.824	5.737	5.650
8	5.814	5.667	5.515	5.359	5.279	5.198	5.116	5.032	4.946	4.859
9	5.257	5.111	4.962	4.808	4.729	4.649	4.567	4.483	4.398	4.311
10	4.849	4.706	4.558	4.405	4.327	4.247	4.165	4.082	3.996	3.909
11	4.539	4.397	4.251	4.099	4.021	3.941	3.860	3.776	3.690	3.602
12	4.296	4.155	4.010	3.858	3.780	3.701	3.619	3.535	3.449	3.361
13	4.100	3.960	3.815	3.665	3.587	3.507	3.425	3.341	3.255	3.165
14	3.939	3.800	3.656	3.505	3.427	3.348	3.266	3.181	3.094	3.004
15	3.805	3.666	3.522	3.372	3.294	3.214	3.132	3.047	2.959	2.868
16	3.691	3.553	3.409	3.259	3.181	3.101	3.018	2.933	2.845	2.753
17	3.593	3.455	3.312	3.162	3.083	3.003	2.920	2.835	2.746	2.653
18	3.508	3.371	3.227	3.077	2.999	2.919	2.835	2.749	2.660	2.566
19	3.434	3.297	3.153	3.003	2.925	2.844	2.761	2.674	2.584	2.489
20	3.368	3.231	3.088	2.938	2.859	2.778	2.695	2.608	2.517	2.421
21	3.310	3.173	3.030	2.880	2.801	2.720	2.636	2.548	2.457	2.360
22	3.258	3.121	2.978	2.827	2.749	2.667	2.583	2.495	2.403	2.305
23	3.211	3.074	2.931	2.780	2.702	2.620	2.536	2.447	2.354	2.256
24	3.168	3.032	2.889	2.738	2.659	2.577	2.492	2.403	2.310	2.211
25	3.129	2.993	2.850	2.699	2.620	2.538	2.453	2.364	2.270	2.169
26	3.094	2.958	2.815	2.664	2.585	2.503	2.417	2.327	2.233	2.131
27	3.062	2.926	2.783	2.632	2.552	2.470	2.384	2.294	2.198	2.097
28	3.032	2.896	2.753	2.602	2.522	2.440	2.354	2.263	2.167	2.064
29	3.005	2.868	2.726	2.574	2.495	2.412	2.325	2.234	2.138	2.034
30	2.979	2.843	2.700	2.549	2.469	2.386	2.299	2.208	2.111	2.006
40	2.801	2.665	2.522	2.369	2.288	2.203	2.114	2.019	1.917	1.805
60	2.632	2.496	2.352	2.198	2.115	2.028	1.936	1.836	1.726	1.601
120	2.472	2.336	2.191	2.035	1.950	1.860	1.763	1.656	1.533	1.381
∞	2.321	2.185	2.039	1.878	1.791	1.696	1.592	1.473	1.325	1.005

Repartiția χ^2

Nr. gradelor de libertate	Probabilitatea $P(\chi^2 > \chi^2_\epsilon)$							
	0.99	0.98	0.95	0.9	0.8	0.7	0.5	0.3
1	0.000	0.001	0.004	0.016	0.064	0.148	0.455	1.074
2	0.020	0.040	0.103	0.211	0.446	0.713	1.386	2.408
3	0.115	0.185	0.352	0.584	1.005	1.424	2.366	3.665
4	0.297	0.429	0.711	1.064	1.649	2.195	3.357	4.878
5	0.554	0.752	1.145	1.610	2.343	3.000	4.351	6.064
6	0.872	1.134	1.635	2.204	3.070	3.828	5.348	7.231
7	1.239	1.564	2.167	2.833	3.822	4.671	6.346	8.383
8	1.646	2.032	2.733	3.490	4.594	5.527	7.344	9.524
9	2.088	2.532	3.325	4.168	5.380	6.393	8.343	10.656
10	2.558	3.059	3.940	4.865	6.179	7.267	9.342	11.781
11	3.054	3.609	4.575	5.578	6.989	8.148	10.341	12.899
12	3.571	4.178	5.226	6.304	7.807	9.034	11.340	14.011
13	4.107	4.765	5.892	7.042	8.634	9.926	12.340	15.119
14	4.660	5.368	6.571	7.790	9.467	10.821	13.339	16.222
15	5.229	5.985	7.261	8.547	10.307	11.721	14.339	17.322
16	5.812	6.614	7.962	9.312	11.152	12.624	15.339	18.418
17	6.408	7.255	8.672	10.085	12.002	13.531	16.338	19.511
18	7.015	7.906	9.390	10.865	12.857	14.440	17.338	20.601
19	7.633	8.567	10.117	11.651	13.716	15.352	18.338	21.689
20	8.260	9.237	10.851	12.443	14.578	16.266	19.337	22.775
21	8.897	9.915	11.591	13.240	15.445	17.182	20.337	23.858
22	9.542	10.600	12.338	14.041	16.314	18.101	21.337	24.939
23	10.196	11.293	13.091	14.848	17.187	19.021	22.337	26.018
24	10.856	11.992	13.848	15.659	18.062	19.943	23.337	27.096
25	11.524	12.697	14.611	16.473	18.940	20.867	24.337	28.172
26	12.198	13.409	15.379	17.292	19.820	21.792	25.336	29.246
27	12.879	14.125	16.151	18.114	20.703	22.719	26.336	30.319
28	13.565	14.847	16.928	18.939	21.588	23.647	27.336	31.391
29	14.256	15.574	17.708	19.768	22.475	24.577	28.336	32.461
30	14.953	16.306	18.493	20.599	23.364	25.508	29.336	33.530

Nr. gradelor de libertate	Probabilitatea $P(\chi^2 > \chi^2_\epsilon)$							
	0.2	0.1	0.05	0.02	0.01	0.005	0.002	0.001
1	1.642	2.706	3.841	5.412	6.635	7.879	9.550	10.828
2	3.219	4.605	5.991	7.824	9.210	10.597	12.429	13.815
3	4.642	6.251	7.815	9.837	11.345	12.838	14.795	16.267
4	5.989	7.779	9.488	11.668	13.277	14.860	16.924	18.467
5	7.289	9.236	11.070	13.388	15.086	16.750	18.908	20.515
6	8.558	10.645	12.592	15.033	16.812	18.548	20.791	22.458
7	9.803	12.017	14.067	16.622	18.475	20.278	22.601	24.321
8	11.030	13.362	15.507	18.168	20.090	21.955	24.352	26.125
9	12.242	14.684	16.919	19.679	21.666	23.589	26.056	27.877
10	13.442	15.987	18.307	21.161	23.209	25.188	27.722	29.588
11	14.631	17.275	19.675	22.618	24.725	26.757	29.354	31.264
12	15.812	18.549	21.026	24.054	26.217	28.299	30.957	32.909
13	16.985	19.812	22.362	25.472	27.688	29.819	32.535	34.529
14	18.151	21.064	23.685	26.873	29.141	31.319	34.091	36.124
15	19.311	22.307	24.996	28.260	30.578	32.801	35.628	37.698
16	20.465	23.542	26.296	29.633	32.000	34.267	37.146	39.253
17	21.615	24.769	27.587	30.995	33.409	35.719	38.649	40.790
18	22.760	25.989	28.869	32.346	34.805	37.156	40.137	42.312
19	23.900	27.204	30.144	33.687	36.191	38.582	41.610	43.821
20	25.037	28.412	31.410	35.020	37.566	39.997	43.072	45.315
21	26.171	29.615	32.671	36.343	38.932	41.401	44.522	46.797
22	27.301	30.813	33.924	37.659	40.289	42.796	45.962	48.268
23	28.429	32.007	35.172	38.968	41.638	44.181	47.392	49.728
24	29.553	33.196	36.415	40.270	42.980	45.559	48.812	51.179
25	30.675	34.382	37.652	41.566	44.314	46.928	50.224	52.620
26	31.795	35.563	38.885	42.856	45.642	48.290	51.627	54.052
27	32.912	36.741	40.113	44.140	46.963	49.645	53.023	55.477
28	34.027	37.916	41.337	45.419	48.278	50.993	54.411	56.892
29	35.139	39.087	42.557	46.693	49.588	52.336	55.792	58.301
30	36.250	40.256	43.773	47.962	50.892	53.672	57.167	59.702



Valorile critice ale repartiției Student cu ν grade de libertate

	$t_{0.1}$	$t_{0.05}$	$t_{0.025}$	$t_{0.01}$	$t_{0.005}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
∞	1.282	1.645	1.960	2.326	2.576

BIBLIOGRAFIE

28. Grotzinger, J.M. - Methods and techniques of research. Ed. Economics Press (1989)
27. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
26. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
25. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
24. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
23. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
22. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
21. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
20. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
19. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
18. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
17. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
16. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
15. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
14. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
13. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
12. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
11. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
10. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
9. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
8. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
7. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
6. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
5. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
4. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
3. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
2. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)
1. Grotzinger, J.M. - The theory of the firm. Ed. Economics Press (1981)

1. Agresti, A. – *Categorical data analysis*, Ed. John Wiley & Sons, 1990
2. Andrei, T. și Stancu, S. – *Statistică – teorie și aplicații* – Ed. All, București, 1995
3. Ardilly, P. – *Les techniques de sondage*, Ed. Technip, Paris, 1994
4. Bărbat, Al. – *Teoria statisticii sociale*, E.D.P. București, 1972.
5. Baron, T.; Biji, E.; Tövissi, L., ș.a. – *Statistică teoretică și economică*, E.D.P. București, 1991.
6. Barițiu, G. – *Scrieri social-politice, studiu și antologie* de V. Chereșteșiu, Ed. Politică, București, 1962.
7. Berdot, J.-P. – *Econometrie*, Université de Poitiers, CNED, 2001.
8. Biji, M.; Biji, El. – *Statistica teoretică*, E.D.P. București, 1979.
9. Biji, M.; Stoichiță, I. – *Metoda selectivă a cercetării statistice*, Ed. Științifică, București, 1957.
10. Calot, G. – *Cours de statistique descriptive*, Dunod, Paris, 1975.
11. Capanu, I; Wagner, P.; Mitruț, C. – *Sistemul conturilor naționale și agregate macroeconomice*, Ed. All, București, 1994.
12. Chelcea, S. – *Chestionarul în investigația sociologică*, Ed. Științifică și Enciclopedică, București, 1975.
13. Ciucu, G.; Craiu, V.; Ștefănescu, V. – *Statistică matematică și cercetări operaționale*, E.D.P., 1978.
14. Constantin, Fl. – *Constantin Mavrocordat*, Ed. Militară, București, 1985
12. Cormier, Ch. – *Les Panels*, în volumul *Probleme actuale de statistică*, Editura Junimea, Iași, 2004
13. Craiu, V. – *Verificarea ipotezelor statistice*, E.D.P., București, 1972.
14. Deroo, M.; Dussaix, A.-M. – „*Pratique et analyse d'enquêtes par sondage*“, Ed. PUF, 1980.
15. Dodge, Y. – *Statistique*, Neuchâtel, P.A.N., 1990
16. Dodge, Y. – *Alternative Methods of Regression*, New York, John Wiley, 1993
17. Driesbeke, J.; Fichet, B.; Tassi, Ph.; éditeurs – *Les sondages*, Ed. Economica, Paris, 1987
18. Driesbeke, J. – *Éléments de statistique*, Ed. Ellipse, Paris, 1992.
19. Dubois, J. – *Méthodologie économique et technique statistique*, Montpellier, 1965
20. Fisher, I. – *The making of index numbers*, New York, 1922.
21. Florea, I. – *Statistică*, Univ. Cluj, 1986.
22. Galton, Fr. – *Natural Inheritance*, Macmillan, London, 1889
23. Georgescu – Roegen, N. – *Metoda statistică – elemente de statistică matematică*, I.S.C.S., București, 1933.
24. Georgescu – Roegen, N. – *Legea entropiei și procesul economic*, Ed. Politică, București, 1979.
25. Giard, V. – *Statistique appliquée à la Gestion*, Ed. Economica, Paris, 1985
26. Grama, A., Băcâin I., Grama, V., Mihalache, S., Genete, D., Sireteanu, A., Trifan, R., Rusu, D., Homocianu, D., Neagu, D., *Sub Windows să învățăm Word, PowerPoint, Front Page și Internet*, Editura Sedcom Libris, Iași, 2004
27. Gourieroux, Ch. – *Théorie des sondages*, Economica, Paris, 1981
28. Grosbras, J.M. – *Méthodes statistiques des sondages*, Ed. Economica, Paris, 1987

29. Guitton, Henri – *Statistique*, Dalles, Paris, 1967.
30. Hașigan, D.O. – *Metodele reprezentării grafice a datelor statistice*, Ed. Științifică, București, 1958
31. Horomnea, E., *Bazele contabilității. Concepte, aplicații, lexicon*, Editura Sedcom Libris, Iași, 2003.
32. Isaic-Maniu, Al.; Grădinaru, A.; Voineagu, V.; Mitruț, C. – *Statistică teoretică și economică*, Ed. Tehnică, Chișinău, 1994.
33. Jaba, Elisabeta; Niculiciou, P.; Bilaus, M. – *Cercetarea selectivă*, Univ. „Al.I. Cuza” Iași, 1977
34. Jaba, Elisabeta – *Statistica. Statistica economiei naționale*, Univ. „Al.I. Cuza” Iași, 1982.
35. Jaba, Elisabeta – *Statistica. Sistem metodologic. Aplicații.*, Univ. „Al.I. Cuza” Iași, 1986.
36. Jaba, Elisabeta; Grama, Ana – *Analiza statistică cu SPSS sub Windows*, Ed. Polirom, Iași, 2004
37. Jaba, Elisabeta – *Statistica*, Ed. Sedcom Libris, Iași, 1996
38. Jaba, Elisabeta – *Statistica*, Ediția a treia, Ed. Economica, București, 2002
39. Jaba, Elisabeta, Pintilescu, Carmen, Iosub Floarea, *Statistică descriptivă. Teste grilă și probleme*, Editura Sedcom Libris, Iași, 2001.
40. Keller, G.; Warraek, B. – *Essentials of Business Statistics*, Wadsworth, Inc., 1991
41. Korka, M., Begu, L., Roșca, E., Tușa E. – *Aplicații statistice*, Editura Bucovina Viitoare, Suceava, 2002
42. Lange, J. – *Éléments de techniques statistique*, Dunod, Paris, 1968.
43. Lecaillon, J.; Labrousse, C. – *Statistique descriptive*, Ed. Cujas, Paris, 1988
44. Mallinvaud, E. – *Méthodes statistiques de l'économetrie*, Dunod, Paris, 1981
45. Marin, D. ș.a. – *Teoria echilibrului general*, Ed. Omnia, UNISAST, Brașov, 1995
46. Marinescu, I. – *Analiza factorială*, Ed. Șt. și Enciclopedică, București, 1984.
47. McClave, J.T.; Benson, P.G. – *Statistics for Business and Economics*, Dellen Publishing Company, San Francisco, 1985.
48. Mihoc, Gh.; Craiu, V. – *Tratat de statistică matematică*, Ed. Acad., București, 1976-1977.
49. Mihoc, Gh.; Urseanu, V. – *Sondaje și estimări statistice. Teorie și aplicații*, Ed. Tehnică, București, 1977
50. Mihoc, Gh.; Urseanu, V.; Ursianu, Em. – *Modele de analiză statistică*, Ed. Șt. și Encicl., București, 1982.
51. Mills, Fr.G. – *Statistical Methods*, 3rd ed., Henry Holt, New York, 1955.
52. Milton Smith, G. – *Ghid simplificat de statistică pentru psihologie și pedagogie*, E.D.P., București 1971.
53. Moineagu, C. – *Modelarea corelațiilor în economie*, Ed. Șt., București, 1974.
54. Moisil, Gr.C. – *Îndoeli și certitudini*, Ed. Enciclopedică Română, București, 1974
55. Morgenstern, Oskar – *Précision et incertitude des données économiques*, Dunod, Paris, 1972.
56. Moser, C.A. – *Metodele de anchetă în investigarea fenomenelor sociale*, Ed. Științifică, Buucurești, 1967.

57. Nenciu, E. – *Probabilități și statistică matematică*, Univ., „Al.I.Cuza” Iași, 1986.
58. Noica, C. – *Jurnal filozofic*, Ed. Humanitas, București, 1990
59. Onicescu, O.; Ștefănescu, V. – *Elemente de statistică informațională cu aplicații*, Ed. Tehnică, București, 1979.
60. Oprea, D., *Analiza și proiectarea sistemelor informaționale economice*, Editura Polirom, Iași, 1999
61. Paraschivescu, M.; Grama, A.; Oprea, D.; Airinei, D. – *Proiectarea sistemelor informatice*, Ed. Univ., „Al.I.Cuza”, Iași, 1989
62. Pecican, E.S. – *Econometrie*, Ed. All, București, 1994
63. Pintilescu, Carmen, *Analiza datelor*, Editura Junimea, Iași, 2002.
64. Porojan, D. – *Statistica și teoria sondajului*, Ed. Șansa, București, 1993
65. Piatier, A. – *Statistique descriptive et initiation à l'analyse*, Themis, Paris, 1962
66. Retegan, G.; Păcurariu, I. coord. – *Populația României*, Ed. Meridiane, București, 1974.
67. Râmneamțu, P. – *Elemente de biometrie medicală și statistică vitală*, Inst. Central de Statistică, București, 1939.
68. Rumșinski, L.Z. – *Prelucrarea matematică a datelor experimentale*, Ed. Tehnică, București, 1974.
69. Saporta, G. – *Probabilités, analyse des données et statistique*, Ed. Technip, Paris, 1990
70. Singly, Francois de – „*L'enquête et ses méthodes: le questionnaire*”, Ed. Nathan, Paris, 1992.
71. Snedecor, G.W. – *Metode statistice aplicate în cercetările din agricultură și biologie*, E.D.P., București, 1968.
72. Snedecor G.W.; Cochran, W.G. – *Statistical Methods*, Yowa State University Press, 1969
73. Sora, V.; Hristache, I.; Mihăescu, C. – *Demografie și statistică socială*, Ed. Economică, București, 1996.
74. Ștrumilin, S.G. – *Scrieri economice alese*, Ed. Științifică, București, 1972
75. Tabără, N. – *Contabilitate națională*, Ed. Moldova, Iași, 1996
76. Tiron, M. – *Teoria erorilor de măsurare și metoda celor mai mici pătrate*, Ed. Tehnică, București, 1972.
77. Tövissi, L.; Vodă, V. – *Metode statistice. Aplicații în producție.*, Ed. Șt. și Encicl., București, 1982.
78. Tövissi, L.; Isaic-Maniu, Al. – *Statistica*, A.S.E., București, 1984.
79. Trebici, V.(coord.) – *Mica enciclopedie de statistică*, Ed. Șt. și Encicl., București, 1985.
80. Țarcă, M. – *Statistică*, vol. I și II, Univ. „Al.I.Cuza” Iași, 1979.
81. Urseanu, V. – *Metode statistice de cercetare a legităților ce formează și mișcare a prețurilor*, Ed. Acad., București, 1973.
82. Yule, U.G.; Kendall, M.C. – *Introducere în teoria statisticii*, Ed. Șt., București, 1969.
83. Zaiț, D., Nica, P. – *Introducere în modelarea econometrică*, Ed. Univ. „Al.I. Cuza” Iași, 1995
84. * * * – *La qualité de l'information dans les enquêtes. Journées d'études de l'ASV*, 1991
85. * * * – *Pour une histoire de la statistique*, Tome II, Paris, 1977

CULOAREA ESTE O PASIUNE, IAR PASIUNEA SE ÎMPARTE...

tipografia



TIPOGRAFIA SEDCOM LIBRIS

IAȘI, ROMÂNIA

ISO 9001:2000

Șos. Moara de Foc nr. 4,

cod 700572, Iași, România

Tel.: +40.232.234.582, 270.479, 242.877;

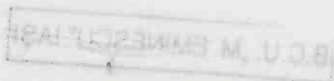
Fax.: +40.232.233.080

E-mail: tipografia@sedcom.ro

<http://www.sedcom.ro>

libris

PENTRU CĂ NE PASĂ CINE EȘTI, CE FACI, CUM FACI ȘI UNDE VREI SĂ AJUNGI...



57. Născu, H. - *Probabilități și statistici matematice*, Univ. „Al.I. Cuza” Iași, 1986.
58. Născu, C. - *Jurnal Biologic*, Ed. Humanitas, Iași, 1990.
59. Olărescu, O.; Șelăreanu, V. - *Elemente de statistică informatică cu aplicații*, Ed. Tehnică, București, 1979.
60. Opres, D.; *Analiza și proiectarea sistemelor informaționale economice*, Editura Pehnam, Iași, 1997.
61. Paraschivescu, M.; Orban, A.; Opres, D.; Airimă, D. - *Proiectarea sistemelor informatice*, Ed. Univ. „Al.I. Cuza”, Iași, 1989.
62. Pecican, E.S. - *Ecumenic*, Ed. All. Doc. univ., 1994.
63. Pinăreanu, Carmen. *Amplasarea și delor*, Editura Junimea, Iași, 2007.
64. Pomian, D. - *Statistica și teoria sondajului*, Ed. Santa, București, 1993.
65. Potier, A. - *Statistique descriptive et initiation à l'analyse*, Tachis, Paris, 1962.
66. Roșanu, G.; Pacurar, I. coord. - *Populația României*, Ed. Meridiane, București, 1978.
67. Rădulescu, P. - *Elemente de biometrie medicală și statistică medicală*, Inst. Central de Statistică, București, 1938.
68. Rădulescu, L.Z. - *Probabilități matematice a datelor experimentale*, Ed. Tehnică, București, 1974.
69. Saporta, G. - *Probabilités, analyse des données et statistique*, Ed. Technip, Paris, 1987.
70. Shaply, Borisovici - *L'enquête et ses méthodes: le questionnaire*, Ed. Nathan, Paris, 1992.
71. Snedecor, G.W. - *Methods of statistical analysis in agricultural and biological*, E.D.P., București, 1968.
72. Snedecor G.W.; Cochran W.G. - *Statistical Methods*, Iowa State University Press, 1969.
73. Soru, V.I.; Hristache, L.; Mărușter, C. - *Demografie și statistică socială*, Ed. Economica, București, 1979.
74. Străduț, S.G. - *Serții matematice algebre*, Ed. Științifică, București, 1972.
75. Tabără, M. - *Contabilitate*, Editura Științifică, București, 1973.
76. Tiron, M. - *Proiectarea și realizarea sistemelor informaționale*, Ed. Tehnică, București, 1979.
77. Tăvăs, L.; Voică, V. - *Statistica socială*, Ed. Științifică, București, 1982.
78. Tăvăs, L.; Voică, V. - *Statistica socială*, Ed. Științifică, București, 1982.
79. Tăvăs, V. coord. - *Microstatistica de calcul*, Ed. Științifică, București, 1985.
80. Tăvăs, M. - *Statistica*, vol. I și II, Univ. „Al.I. Cuza” Iași, 1978.
81. Ursanu, V. - *Metode statistice de cercetare în biologie și medicină*, Ed. Științifică, București, 1973.
82. Yule, U.; Kendall, M.C. - *Introducere în teoria statistică*, Ed. Șt., București, 1969.
83. Zait, P. - *Introducere în modelarea econometrică*, Ed. Univ. „Al.I. Cuza” Iași, 1991.
84. - *La qualité de l'information dans les enquêtes*, Journales d'études de l'IASI, 1991.
85. - *Pour une théorie de la statistique*, Journales d'études de l'IASI, 1991.

BCU IASI/CENTRAL UNIVERSITY LIBRARY

BCU IASI/CENTRAL UNIVERSITY LIBRARY

I.S.B.N.: 978-973-670-227-3



6 420635 202398